Essays on Pension Finance and Dynamic Asset Allocation

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof.dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op maandag 21 juni 2010 om 14.15 uur door

Renxiang Dai

geboren op 11 juni 1974 te Anhua, Hunan, China.

PROMOTOR: prof.dr. J.M. Schumacher

To my family.

Acknowledgements

This thesis benefited enormously from many people's help and support in the three years when this project was carried out. I would like to express my gratitude to some of them, although the list of individuals I wish to thank would be beyond what this short text could accommodate.

First of all, I thank Prof. Hans Schumacher for his enthusiastic supervision. Without his guidance and support, the thesis as is presented here could not have been accomplished. What I have learnt from him, I believe, will be a valuable asset in my life.

I am grateful to the thesis committee members, Prof. Joost Driessen, Prof. Theo Nijman, Dr. Juan Carlos Rodriguez, and Prof. Bas Werker, for their helpful comments and suggestions. Their insight and wisdom helped greatly to improve upon the thesis.

This PhD project was supported by Netspar; I gratefully acknowledge its funding and highly motivating research environment. During this project, I worked at the Department of Econometrics and Operations Research, and I would like to thank all my colleagues for this pleasant experience. The research facilities and support provided by the CentER Graduate School and the CentER Finance Group are acknowledged with gratitude.

I would like to thank my friends and fellow PhD students at Tilburg University for many memorable times we had together. They have helped greatly to make the three years in the Netherlands enjoyable for my family and me. It is impossible to list their names without omission, so I save the efforts. My deep gratitude to them, however, cannot be overestimated.

I am indebted to my parents and my extended family for their understanding and support. The vast distance between us in those years was unable to stop me from receiving and feeling their care and love.

To my wife, Yan, and my son, Molei, I would like to say thank you so much. Your company made the years in Tilburg much more pleasant, and your understanding and love are always a source of inspiration to me.

April, 2010

Renxiang Dai

Contents

Chapter 1

Introduction

In most countries, pensions are provided in two major forms: defined benefit (DB) and defined contribution (DC). Under DB pension schemes, the employee's pension benefit is determined by a formula that takes into account such factors as years of service for the employer and in most cases, wages or salary. Pension legislation often requires plan sponsors to make good on these promises even if the underlying value of the pension reserve falls short. Thus, pension sponsors, rather than pension plan participants, bear pension investment and longevity risks. The aging population and the international move towards the market-based accounting standard, however, have placed substantial funding pressure on DB plans. As a consequence, the past two decades have seen a strong trend away from DB plans toward defined-contribution (DC) plans in many countries. Under a typical DC plan, each participant has an individual retirement account into which the participant and the sponsors (if any) make regular contributions. The retirement benefit then depends on the total contribution and investment earnings of the accumulation in the account over time. In the case of DC plans, retirement saving and income tend to be more subject to employees' control throughout the life cycle, and hence it helps to relieve employers and other sponsors of some, if not all, responsibility for pension provision under the DB framework.

Many commentators seem to agree that the shift to DC plans, however, is far from being a satisfactory solution because they are too complex and too risky for individuals. Individuals typically lack the financial expertise and computation capacities to implement complex lifetime financial planning. DC plans are also vulnerable to large marketing and management costs, and to market failure like that stemming from adverse selection in annuity markets. As a balance between DB and DC, an approach which has been implemented recently by many Dutch pension funds and which is under discussion in the UK is to introduce a practice known as "conditional indexation".

The way in which a conditional indexation scheme operates can be best illustrated by an example. Let the financial position of a pension fund be measured by its *funding ratio*, i. e. the ratio of pension asset value to pension liability value. The level of compensation for inflation (indexation) applied in a certain year is determined according to a rule of the following form: (i) if the funding ratio is below some threshold (e. g. 110%), there is no indexation to inflation; (ii) if the funding ratio is above some upper threshold (e. g. 140%), the pension rights will be fully indexed to inflation; (iii) if the ratio is in between, some intermediate level of indexation will apply. The rule that links the funding ratio to the indexation decision is referred to as a "policy ladder". Several large pension funds in the Netherlands have published such policy ladders.

From the perspective of participants, conditional indexation schemes are similar to DB plans in the way benefits are specified. From the perspective of pension funds, conditional indexation brings a DC element as the liability value will generally change in line with the development of the fund's asset value through the practice of indexation, therefore serving as a shield of the funding ratio against the fluctuation of asset value stemming from exposure to financial markets. In a nutshell, conditional indexation, like traditional DB, enables participants to enjoy a high level of pension predictability, and like DC, enables pension funds to have a high level of financial stability.

The introduction of conditional indexation may raise a number of issues, and in this dissertation, Chapters 2 and 3 are devoted to address some of them. In Chapter 2, we focus on the quality of pension profile of conditional indexation schemes from a life-cycle investment perspective, working under the assumption that participation in the scheme is compulsory (as is the case in the Netherlands) so that in practice a financial constraint is imposed on participants. Welfare analysis is applied to investigate the performance of such schemes relative to alternative investment strategies such as fixed-mix policies. We carry out this analysis in the context of a broad family of utility functions, which takes into account the possible presence of two benchmark levels corresponding to a minimum guaranty and to full indexation respectively. For the purpose of comparability, we construct a self-financing continuous-time implementation of the conditional indexation scheme. The implementation involves continual adjustment of the parameters of the contingent claim representing final payoff. Our findings indicate that, in situations where large weight is placed on the benchmark levels, conditional indexation is fairly close to being optimal.

Against the backdrop of the international move promoted by the International Accounting Standards Board towards the market-based, fair value accountancy standard, the valuation of pension liabilities with conditional indexation have become a subject of increasing interest both to academics and practitioners. In Chapter 3, we consider the valuation of conditionally indexed pension liabilities in the framework of market valuation. We formulate the circularity problem that may arise in the valuation of conditionally indexed pension liabilities. Namely, the funding ratio determines the indexation level through a chosen indexation rule (often known as a "policy ladder"), but at the same time the indexation level may, with market valuation of pension liabilities, have a feedback effect on the liability value and in turn on the funding ratio. We develop a backward recursion approach to the valuation of liabilities subject to the circularity constraint. Numerical examples are used to show the impact of investment strategies and indexation rules on the liability value. The current implementation of conditional indexation uses as the basis for indexation decisions a proxy of funding ratio, rather than the funding ratio based on market-based valuation, and in this way avoids the circularity problem. Our findings show that the proxy of funding ratio may be misleading in assessing the actual financial status of pension funds, and for this purpose the actual funding ratio needs to be computed and used.

Another theme addressed in the thesis is asset allocation. Aside from its relevance for pension fund management, how to invest and consume in the best way has broader implications for individuals and institutional investors. In this field, there are many problems yet to be solved, and the other two chapters of the thesis address two of them.

There has been well documented empirical evidence that the equity premia are time-varying rather than constant. Expected returns on common stocks have been found to vary over business cycles, and the general message of the empirical evidence is that expected returns are lower when business conditions are strong and higher when business conditions are weak. In this situation, optimal consumption and investment strategies involve timing business conditions. On the other hand, it has long been recognized that investors do not have complete information about business conditions or the state of economy. Institutional and individual investors usually disagree on the assessment of the current business conditions, and economic prospects. This disagreement is symptomatic of the fact that there is quite some uncertainty surrounding the state of the economy. In the context of time-varying expected returns, state uncertainty leads to uncertainty about expected returns, which an investor must take into account in choosing the optimal consumption and investment.

In Chapter 4, we examine the optimal consumption and asset allocation in a setting where expected returns on stocks are time-varying, but unobservable. The dynamic optimization problem is addressed in two separate steps, namely estimation and optimization. The optimal consumption and investment plans can be expressed in closed form in the stable phase where the investor has a sufficiently long history of the stock price that she can no longer improve upon the estimation error. Numerical exercises show that there is significant market timing in the optimal strategy. We also discuss the implications of this study for the benefit policy of pension funds, and find that the benefit should be relatively stable over time when a constant-proportion investment strategy is employed.

Chapter 5 is devoted to the question how to include the asset class of commodities into the traditional portfolio of stocks and bonds. Commodities have been emerging as an increasingly important class of assets for institutional and individual investors in recent years. Around 2007, the size of the global commodities derivatives market is estimated to be about 750 billion US dollars. As the markets have grown, more investors have been attracted to commodities. Increased exposure to commodities has been acquired by both institutional and individual investors.

Chapter 5 studies commodity investment in the context of dynamic asset allocation, with a focus on the implications of the commodity return predictability arising from mean reverting commodity prices. The model of financial markets consists of three asset classes: stocks, bonds, and commodities, which generalizes the benchmark setting of Merton (1969). The risk premium in the commodity market is assumed to be dependent on the mean-reverting spot commodity price, and this assumption is supported by the empirical findings of the paper. I solve, in closed form, the optimal portfolio and consumption strategies. The study suggests that allocation to commodities is needed to optimize the instantaneous risk-return profile (myopic purposes), as well as to hedge the stochastic changes of the investment opportunity set (intertemporal purposes). The welfare cost of excluding the commodity from financial decision making is also solved in closed form. A simple numerical exercise shows that there is substantial market timing in the optimal financial policy, and that excluding the asset class of commodities may incur substantial welfare costs, especially for long-term and less risk-averse investors.

The thesis is an outcome of research cooperation: Chapters 2 and 3 are joint work with J.M. Schumacher, whereas Chapter 4 is coauthored with J.C. Rodriguez and J.M. Schumacher.

Chapter 2

Welfare Analysis of Conditional Indexation Schemes from a Two-Reference-Point Perspective

2.1 Introduction

Over the years, employers around the world have strived to provide retirement income security by setting up defined-benefit (DB) pension schemes. Under such schemes, the employee's pension benefit is determined by a formula that takes into account such factors as years of service for the employer and wages or salary, and it is pension sponsors, rather than pension plan participants, who bear pension investment and longevity risks. For pension participants, an important appeal of the DB model is that it allows them to plan their retirement income without requiring much knowledge about saving, portfolio choice, capital market risks, or mortality trends. DB plans, however, have been faced with substantial funding pressure in the past two decades because of the aging population and the move towards the market-based accounting standard. And there has been a strong trend away from DB plans toward defined-contribution (DC) plans in many countries, especially in the Anglo-Saxon world. Under a typical DC plan, retirement saving and income tend to be more subject to employees' control throughout the life cycle, and hence it helps to free employers and other sponsors of some, if not all, responsibility for pension provision under the DB framework.

As stated in the Introduction, it has been increasingly recognized that the shift to DC plans is far from being a satisfactory solution because they are too complex and too risky for individuals (see, for instance, Merton [2006]). As is shown by Lusardi and Mitchell [2006] and van Rooij, Lusardi, and Alessie [2006], individuals often lack the financial expertise and computation capacities to carry out complex lifetime financial planning. DC plans are also vulnerable to large marketing and management costs, and to market failure like that stemming from adverse selection in annuity markets. As a balance between DB and DC, an approach which has been implemented recently by many Dutch pension funds is to introduce a practice known as "conditional indexation".

In a conditional indexation scheme, the pension profile of a participant guarantees a *minimum* level which is updated each year through a decision on the inflation indexation for that year on the basis of the funding ratio of pension fund (the ratio of asset value to liability value). That is, the guaranteed level is built up by multiplying a conditionally granted indexation level each year. If a participant is granted full indexation every year, then her pension can fully compensate for inflation, and it is the *maximum* pension she can receive.

Thus conditional indexation schemes are essentially formulated in a framework of two reference points: they guarantee a minimum nominal amount of pension rights, and at the same time aim to provide pension rights sufficient to fully cover inflation. In the words of Bikker and Vlaar [2007],

> "the typical pension contract nowadays comprises an average earnings defined benefit pension in which only nominal benefits are guaranteed, but with the intention to provide wage indexation."

A way of thinking about pension systems is suggested here which has as salient features the presence of both a minimum benefit (guaranteed amount) and a maximum benefit (full indexation).

As a middle way between DB and DC, conditional indexation, like traditional DB, enables participants to enjoy a high level of pension predictability, and like DC, enables pension funds to have a high level of financial stability. While we focus on conditional indexation from a pension perspective in this chapter, similar schemes are also relevant in the context of with-profit policies; cf. for instance Grosen and Jørgensen [2002], Ballotta et al. [2006], and Gatzert and Kling [2007] for a discussion of related schemes.

The introduction of conditional indexation may raise a number of issues, such as the valuation of pension liabilities, the definition of accrued pension rights, and long-term stationarity of pension funds. In this chapter, we focus on the quality of pension profile of conditional indexation schemes from a life-cycle investment perspective, working under the assumption that participation in the scheme is compulsory (as is the case in the Netherlands) so that in practice a financial constraint is imposed on participants.

The objective of carrying out an evaluation of alternative investment schemes calls for the formulation of an evaluation criterion. As is common in the literature, we will use the expected utility framework of von Neumann and Morgenstern [1944]. This framework still allows considerable freedom in choosing a utility function. Rather than summarily eliminating most of this freedom by restricting ourselves to a one-parameter family of utility functions, we apply some considerations relating to the particular nature of the investment scheme that is under investigation in this chapter.

Reference points are absent from the power or constant-relative-risk-aversion (CRRA) utility function, which is the standard criterion underlying much analysis on optimal pension investment and pension scheme design. It is common though for individuals to use benchmarks or reference points as an aid in evaluation and decision-making under uncertainty [Tversky and Kahneman, 1981]. Perhaps the most well-known example to economists is the notion of loss aversion in prospect theory; one of the defining properties of loss aversion is that wealth is measured relative to a given reference point. People divide risky outcomes into gains (greater than the reference point) and losses (less than the reference point), and experiments have shown that people's preferences with respect to gains and with respect to losses are different [Kahneman and Tversky, 1979, Tversky and Kahneman, 1992]. Prospect theory formulates the phenomenon by a utility function with a kink at the reference point.

People may use different reference frameworks for decisions in different situations, and in some cases like the above-mentioned pension fund context it seems more appropriate to use more than one reference point. March and Shapira [1987] argue that two reference points may have significantly more descriptive power than a single one, and that from managerial perspectives on risk taking, a target level for performance and a survival level are the most frequently mentioned references. On the basis of their and other theoretical studies on multiple reference points, Sullivan and Kida [1995] conduct some experiments to investigate the effect of multiple reference points on managers' decision-making under risk. Their finding indicates that presence of two reference points is in conformity with a complex pattern of risk-taking behavior; managers' decisions are affected by the positions of risky alternatives relative to two important reference points.

In the finance literature, the notion of loss aversion, built on the assumption of a single reference point, has been introduced by a number of recent papers, for example, Benartzi and Thaler [1995], Barberis, Huang, and Santos [2001], Berkelaar, Kouwenberg, and Post [2004], and Gomes [2005]. The common modeling approach is to introduce a kink in the reference point distinguishing gains from losses. It may be noted that many commercial investment products involve guarantees; if such products are to be explained as optimal from an expected utility perspective, then kinked utility must play a role. Explanatory factors might include the principal/agent relationship; this factor also plays a role in the pension context. For the purposes of this chapter we assume that all considerations can be sufficiently expressed by a family of kinked utility functions.

Given the salience of two reference points in the pension context, it may be appropriate to allow the presence of two reference points in formulating preferences. The utility function we present below is an extended version of the power utility function, allowing for kinks at two reference points. It is an extension of the power utility function with a kink at a single reference point that Berkelaar et al. [2004] use to express loss aversion.

In addition to reference points, updating the guaranteed level over time as seen in conditional indexation schemes may be justified in part by external habit formation. That is, the reference point that people use to evaluate their consumption depends on the history of general consumption level, reflecting people's desire to "catch up with the Joneses". In this chapter, however, we abstract from this point in formulating a benchmark utility function, in view of the absence of a firmly established standard for expressing external habit formation, and also to stay close to the classical CRRA framework.

The implication of two reference points for pension finance can be shown through looking at the investment policy optimal with respect to the extended power utility function incorporating reference points. Assuming the standard Black and Scholes [1973] economy, we shall solve for the optimal investment policy in the sense that the expected utility of participants is maximized. As will be presented below, the optimal investment strategy can be characterized by a (partial) floor protection at the lower reference point, as well as a (partial) cap at the upper reference point. Intuitively, it can be interpreted as buying partial floor protection through selling part of the upside potential, similarly as in a collar construction. Compared with loss-averse preference which is characterized by utility function with a kink at a single reference point, the strategy optimal for the preference with two reference points provides a better downside protection at the cost of forgoing more upside potential.

The subject of welfare analysis is a stylized conditional indexation scheme which is constructed to have a dynamically updated guaranteed level as seen in practice. To make the welfare analysis comparable to standard life-cycle investment studies, we impose that conditional indexation schemes be financially fair in the sense that the value of pension rights is equal to the value of contributions. In the absence of financial fairness, some participants could be arbitrarily better off with *ex ante* wealth transfer from others. To this end, we shall discuss ways in which one may construct pension systems that combine conditional indexation with financial fairness. The main idea we use below is contingent exchange of one option by another of equal value, with continuous updating of the parameters characterizing the options. This implies that conditional indexation as defined here could in principle be used by an individual as a private investment scheme. Being implemented in a financially fair manner, conditional indexation schemes can then be subject to welfare analysis to see how good conditional indexation schemes can be from the life-cycle investment perspective of participants. We illustrate by numerical examples how the evaluation outcome depends on the presence and strength of the reference points.

Since our purpose in this chapter is to focus on the welfare implications of conditional indexation, we avoid technical complications due to factors that we believe are less directly related to the conditional indexation idea. We do introduce, as discussed above, utility functions that involve particular reference points because such reference points also play a role in conditional indexation. However we do not include in the analysis several features that are often considered in the recent lifecycle investment literature, such as human capital, stochastic interest rates, stochastic inflation, longevity risk, and asset return predictability. In a more comprehensive investigation, such factors should be taken into account; here our aim is to present a first analysis.

The chapter is organized as follows. In the next section, we formulate a class of piecewise power utility functions to allow that risky outcomes are evaluated against two reference points, and specify the financial setting. To establish the benchmark of the welfare analysis, the pension investment optimal for this class of utility functions is investigated in Section 2.3. Section 2.4 discusses the formulation of conditional indexation schemes whose payoff structure is close to those generated by collective pension funds in practice. The welfare analysis of conditional indexation schemes is illustrated by numerical examples in Section 2.5. Some concluding remarks are in Section 5.6.

2.2 The model

2.2.1 The utility function

The power utility function is the most widely used evaluation measure in the literature on dynamic asset allocation. It has some desirable properties in terms of mathematical tractability, and it reflects constant relative risk aversion, which is thought to be a reasonable assumption on people's risk preferences. The power utility function is also used as a building block to accommodate other attributes of preferences like loss aversion and habit formation. In this respect, refer to Berkelaar et al. [2004] for an example on loss aversion, and to Sundaresan [1989] and Campbell and Cochrane [1999] on habit formation. Following a similar approach, we propose a utility function with the power utility as a building block in order to reflect the presence of reference points.

Assume that an individual considers pension investment in a framework of two reference points: a guaranteed level and an intention level, denoted by θ_1 and θ_2 respectively. With respect to the two reference points, possible pension payoffs at retirement, W, can be divided into three regions: below the guaranteed level, beyond the intention level, and in between. As in papers on loss aversion, the presence of the two reference points is formulated by two kinks corresponding to the two points in the utility function. Within each of the three regions, the utility is specified in the standard power form, reflecting *locally* constant relative risk aversion. In addition, we impose that the utility function should be continuous.

In general, the utility function can be formulated by a piecewise power function characterized by five parameters:

$$U(W) = \begin{cases} \kappa_L \theta_1^{\gamma_L - \gamma_M} \big(\psi(W, \gamma_L) - \psi(\theta_1, \gamma_L)\big) + \psi(\theta_1, \gamma_M) & \text{for } W \leq \theta_1, \\ \psi(W, \gamma_M) & \text{for } \theta_1 < W < \theta_2, \\ \frac{1}{\kappa_U \theta_2^{\gamma_M - \gamma_R}} \big(\psi(W, \gamma_R) - \psi(\theta_2, \gamma_R)\big) + \psi(\theta_2, \gamma_M) & \text{for } W \geq \theta_2. \end{cases} \tag{2.1a}$$

where

$$\psi(x, \gamma) = \begin{cases} \frac{x^{1-\gamma} - 1}{1-\gamma} & \text{for } \gamma \neq 1, \\ \log x & \text{for } \gamma = 1. \end{cases} \tag{2.1b}$$

In the above, the parameters γ_L, γ_M, and γ_R are positive and represent the local constant relative risk aversion within the left, middle, and right regions, respectively. The parameters κ_L and κ_U, which must be greater than or equal to 1 to ensure concavity, denote the "kinkedness" of the utility function at the lower and upper reference points.

For simplicity, we mainly work with the two-parameter family that is obtained by the simplification that the degrees of kinkedness are identical at both reference points and the local rates of risk aversion are identical for the three regions:

$$U(W) = \begin{cases} \kappa\psi(W,\gamma) + (1-\kappa)\psi(\theta_1,\gamma) & \text{for } W \leq \theta_1, \\ \psi(W,\gamma) & \text{for } \theta_1 < W < \theta_2, \\ \frac{1}{\kappa}\psi(W,\gamma) + (1-\frac{1}{\kappa})\psi(\theta_2,\gamma) & \text{for } W \geq \theta_2, \end{cases} \tag{2.2}$$

where γ (> 0) is the *local* rate of relative risk aversion, and κ (≥ 1) is the kinkedness parameter (Figure 2.1).

There is a parameter similar to κ in the utility function of prospect theory that represents the degree of loss aversion. Tversky and Kahneman [1992] estimate that the loss aversion parameter is equal to 2.25 based on the experimental results of a group of individuals facing hypothetical decision problems. The two kinks cause the marginal utility to jump at the two reference points. We note that when $\kappa = 1$, the utility function is reduced to the standard power utility function, and that it also incorporates the utility function considered by Berkelaar et al. [2004] as a special case when θ_2 is infinity. Because of the kinks, the preference expressed by the piecewise utility function has the property of first-order risk aversion at the reference points [Segal and Spivak, 1990].

2.2.2 The financial setting

We assume that the individual, over the working life, contributes to an occupational pension scheme an amount whose value is known at entry into the pension system, and receives a lump-sum pension at retirement. In line with standard life-cycle investment analysis, the individual, within the expected utility framework, would invest the contributed amount in such a way that the expected utility over the lump-sum pension is maximized. We work in a highly simplified setting, namely the standard Black-Scholes economy. This assumption, in addition to simplifying the analysis, allows one to focus on the impact on pension investment of two reference points, and makes it straightforward to examine some popular investment policies that already developed in a complete-market setting from a new perspective. Specifically, the financial setting is as follows.

- The only risk factor is stock market risk, and it is traded through a stock index S_t following geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dZ_t,$$

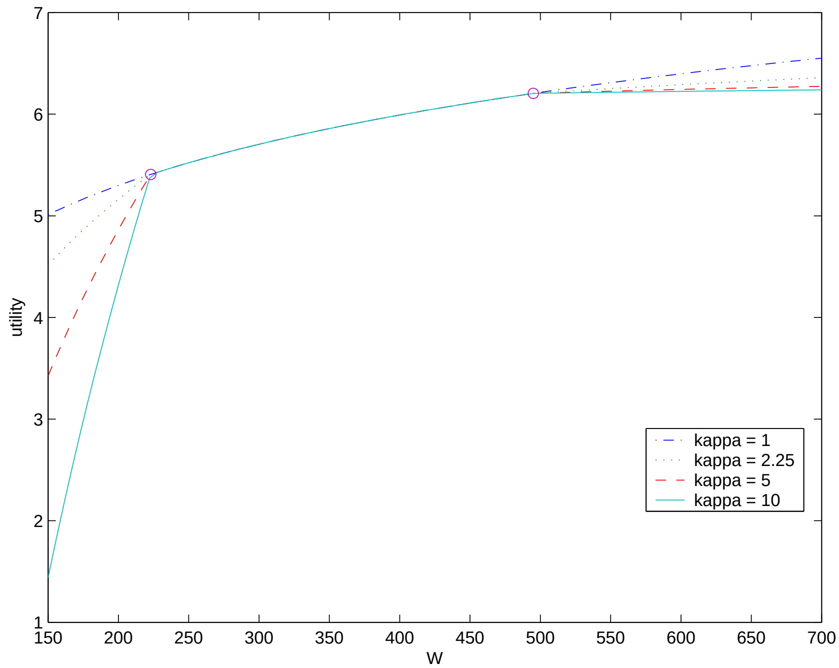

Figure 2.1: **The two-parameter class of utility functions (2.2)** This figure visualizes the utility function for different values of the kinkedness parameter: $\kappa = 1$ (dash-dotted), $\kappa = 2.25$ (dotted), $\kappa = 5$ (dashed), $\kappa = 10$ (solid). Other parameter values are: $\gamma = 1$, $\theta_1 = 223$, and $\theta_2 = 495$.

where μ and σ are the constant drift and volatility parameters, and Z_t is a standard Brownian motion.

- The riskless asset is a cash bond with constant interest rate r, whose price B_t changes according to

$$dB_t = rB_t dt.$$

- Thus the pricing kernel (stochastic discount factor) ξ_t is characterized by

$$d\xi_t = -r\xi_t dt - \lambda\xi_t dZ_t,$$

where $\lambda = \frac{\mu - r}{\sigma}$ is the market price of risk.

- We assume that the individual makes a single contribution at time $t = 0$, and retires and receives pension at $t = T$. The value at time 0 of the contribution is denoted by W_0.

- We assume that two levels θ_1 and θ_2 have been defined which satisfy $\theta_1 < e^{rT} W_0 < \theta_2$ and which are referred to as the *guaranteed level* and the *intention level*, respectively. The corresponding annualized growth rates $\pi_1 := \frac{1}{T}\log(\theta_1/W_0)$ and

$\pi_2 := \frac{1}{T}\log(\theta_2/W_0)$ will for concreteness be referred to as *price inflation* and *wage inflation* respectively. The theory allows other interpretations as well, as long as the inequalities $\pi_1 < r < \pi_2$ are satisfied; for instance π_1 might correspond to a nominal guarantee.

- We shall illustrate results by numerical examples. For this purpose, it is assumed that the economy is characterized by an annual risk-free interest rate of 3%, stock risk premium of 4 percent per year (i. e. $\mu = 7\%$), stock market volatility of 20 percent per year, an annual price inflation of 2 percent and an annual wage inflation of 4 percent. The working life of the individual is 40 years. The present value of the contribution at time $t = 0$ is 100. Thus the guaranteed and intention levels are 223 and 495, respectively. The parameter values are summarized in Table 2.1.

r	μ	σ	π_1	π_2	W_0	T	θ_1	θ_2
3%	7%	20%	2%	4%	100	40	223	495

Table 2.1: **The parameter values**

2.3 Pension investment for the benchmark utility

To understand the benchmark utility function in the context of life-cycle investment, we now investigate the investment policy sought by the individual to optimize $E[U(W_T)]$. In a complete market, such as the Black-Scholes economy, the optimal pension payoff as a function of the state of the economy can be obtained using the equivalent martingale method [Cox and Huang, 1989], i. e.

$$W_T = (U')^{-1}(y\xi_T), \tag{2.3}$$

where $(U')^{-1}$ denotes the inverse of the marginal utility function, and y is a Lagrange multiplier which is determined by the budget constraint $E[\xi_T W_T] = W_0$. In the context of a collective pension fund, the budget constraint can be interpreted as imposing financial fairness between generations.

For the piecewise power utility function (2.2), one can solve the optimal profile of pension W_T as a function of the value of pricing kernel at retirement ξ_T (see Appendix 2.A). To make the optimal pension profile intuitively more appealing, Figure 2.2 plots the optimal pension as a function of the annualized return of stock markets. The optimal

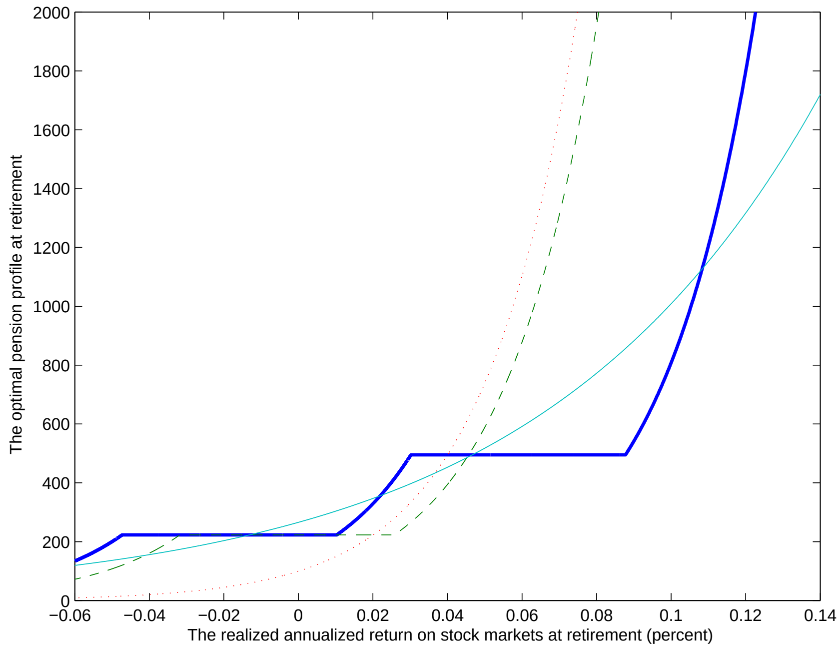

Figure 2.2: **The optimal pension as a function of stock market return** This figure shows the optimal pension payoff for the utility function (2.2) as a function of the stock index return assuming the parameter values: $\kappa = 10$, $\gamma = 1$, $\theta_1 = 223$, and $\theta_2 = 495$ (bold line). For comparison, the figure also shows the optimal profiles for a CRRA utility ($\kappa = 1$, $\gamma = 3$; drawn line), the standard logarithmic utility ($\kappa = 1$, $\gamma = 1$; dotted line), and a utility function with only one kink ($\theta_2 = \infty$; dash-dotted line). For parameter values not mentioned here, see Table 2.1.

pension profile falls into five regions: three slopes connected by two plateaus. In the slope regions, the pension right is increasing with the return on the stock index. In the plateau regions, the pension benefit is constant at the reference levels, independent of stock index changes. The optimal profile of pension benefits can be characterized as a partial floor protection, attained at the cost of forgoing some upside potential of stock markets.

The implication of the two reference points can also manifest itself through comparison to the optimal payoffs for alternative utility functions. Figure 2.2 also shows the payoffs optimal for the standard power preference and the loss-aversion preference. Berkelaar et al. [2004] show that loss aversion as expressed by a single reference point leads to a partial portfolio insurance strategy. In comparison with the loss-averse agent and the CRRA agent, the participant who uses two reference points gives up more payoff in good states of the financial market to finance a better downside protection.

To illustrate the impact on the optimal pension benefit of the two parameters used in (2.2), κ and γ, Figure 2.3 presents the cumulative distribution of the optimal profile of pension benefits for various values of the parameters. The six plots on the first row, assuming $\gamma = 1$, show the simulated cumulative distribution functions of the optimal pension benefits for varying values of κ. For κ greater than 1, the optimal profiles of pension benefits invariably feature probability clustering at the two reference levels, reflecting a partial portfolio insurance and selling of upside potentials. The clustering becomes more pronounced with increasing prominence of the reference points. For the level of kinkedness equal to the rate of loss aversion reported by Tversky and Kahneman [1992], the probabilities of having pension benefits at the guaranteed and intention levels are about 9 percent and 25 percent respectively. For $\kappa = 10$, the clustering becomes dominant, with the probabilities increased to about 10 per cent and 60 percent respectively. If the effect of the reference points is so strong as to justify $\kappa = 100$, then the optimal pension benefit is close to a *binary* payoff structure: the guaranteed level will be paid if the stock market index is below a certain level at retirement, otherwise the intention level will be paid. Actually, it can be shown that a binary payoff structure is optimal in the extreme case where $\gamma = 0$ and $\kappa = \infty$ (See Appendix 2.A).

The plots on the second row of Figure 2.3, assuming $\kappa = 1$, present the distributions of the optimal pension benefits for varying degrees of risk aversion. In this case, the preference reduces to the standard CRRA. As discovered by Samuelson [1969] and Merton [1969], the CRRA individual finds it optimal for allocate a constant proportion of pension asset value in risky assets. This type of strategies, known as "constant-proportion" or "fixed-mix" strategies, lead to a lognormally distributed profile of pension benefit in an economy of the Black-Scholes type. In a given financial market, the proportion in risky assets is determined by the rate of relative risk aversion: in the Black-Scholes economy, the proportion of pension assets in risky assets is $\lambda/\sigma\gamma$, where λ is the market price of risk as introduced before. As shown in Figure 2.3, the variability of the lognormally distributed pension benefits generated by this strategy is decreasing with the degree of risk aversion. For CRRA preference, an increasing degree of risk aversion will make pension benefits concentrate more and more on a single value, rather than on the *two* reference levels as in the case of kinked utility. The difference highlights that κ and γ in the utility function (2.2) reflect different aspects of preference.

To complete the discussion on the investment policy for the benchmark class of utility functions, we turn to the investment strategy needed to realize the optimal profile of pension rights at retirement. The optimal pension can be viewed as a *contingent payoff* that can be replicated by a *delta replication* strategy. The fundamental theorem of asset pricing tells us that the process $\{\xi_t W_t\}$ is a martingale, so the optimal pension asset value

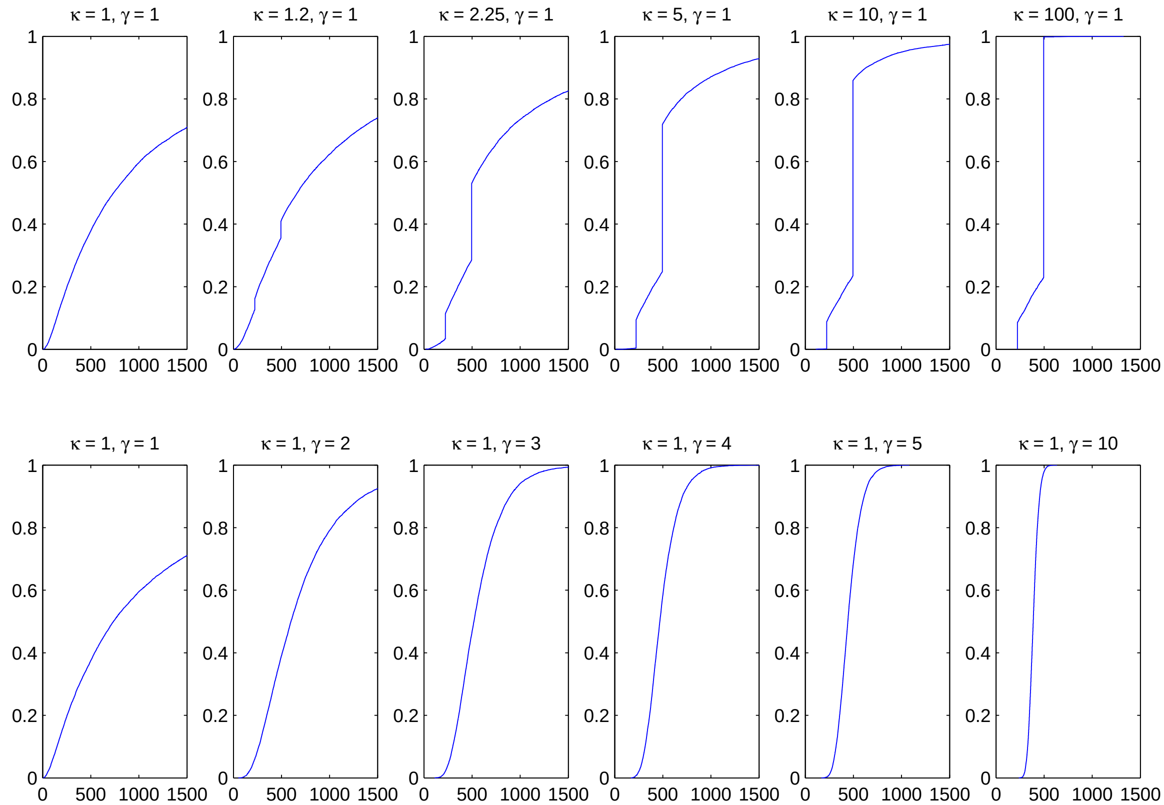

Figure 2.3: **Cumulative distribution of the optimal pension for different values of κ and γ** The figure illustrates the impact of κ and γ on the optimal investment through simulated cumulative distribution function of the resulting pension value at time T. The plots on the first row are for fixed γ and varying values of κ, where those on the second row are for $\kappa = 1$ and varying values of γ. For parameter values not mentioned here, see Table 2.1.

W_t at time t ($\leq T$) satisfies

$$W_t = \frac{1}{\xi_t} E_t[\xi_T W_T]. \tag{2.4}$$

Following this approach, one can solve the optimal pension asset value W_t as a function of time t and the value of the pricing kernel ($W_t = f(t, \xi_t)$).

Given the one-to-one correspondence between the pricing kernel and the stock price in the Black-Scholes economy, the optimal pension asset value can also be expressed as a function of time and of the stock index value S_t, i. e. $W_t = g(t, S_t)$. The holdings of risky assets ("delta") can be determined by the partial derivative of the optimal pension asset value at time t with respect to the stock market level at time t. As an alternative to computing the delta, we characterize the optimal investment strategy by the weight of pension asset value invested in the stock index (see Appendix 2.A). Figure 2.4 illustrates the optimal investment policy by presenting the stock weight as a function of time and the stock market return (only the weights for the last 20 years are shown for ease of

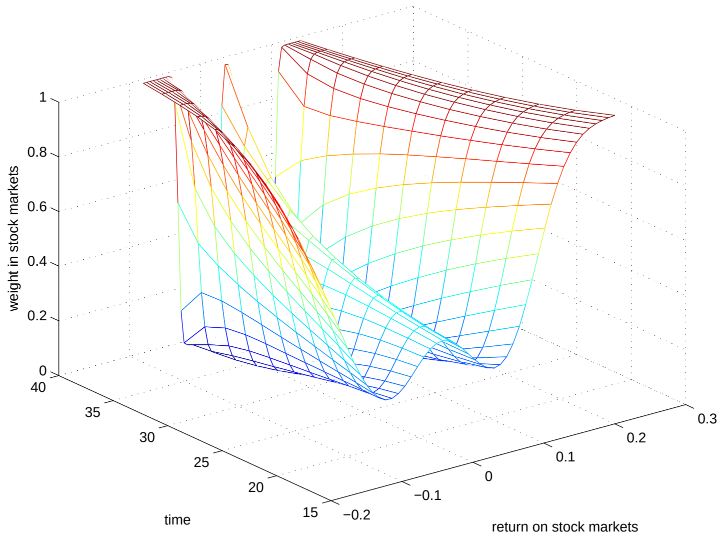

Figure 2.4: **The optimal stock weight** The figure shows the optimal stock weight as a function of time and return on stock markets. The "time" axis represents the calender time, whereas the "return on stock market" axis the return on the stock index so far. It is assumed that $\kappa = 10$ and $\gamma = 1$. For parameter values not mentioned here, see Table 2.1.

viewing). The investment strategy requires a sophisticated, dynamic adjustment of the stock weight depending on time and on the realized return on stock markets. The relation between the weight and the return on stock markets is of a "W" shape. The intuition is as follows. When the return on stock markets is at such levels that it is likely to realize the guaranteed or intention levels, then a low weight in risky assets is needed to ascertain the realization of the reference levels. However, if it is very unlikely for the terminal pension asset value to be at the two reference levels due to, for instance, very strong or weak stock markets, then the pension fund will behave like a constant-relative-risk-averse investor with no kinks, and the weight in risky assets is approaching that required by a constant proportion strategy.[1] This effect is in particular strong at times close to maturity. In scenarios where stock returns are good for some time but then go down, the optimal strategy reduces the stock holdings when bad returns appear, so as to reach at least the upper threshold with high probability.

[1]If there are no kinks ($\kappa = 1$), the resulting CRRA utility with unit rate of relative risk aversion will, for the parameter values considered in the chapter, lead to a constant proportion strategy which allocates 100 percent of the pension asset value in stock markets independent of time and of the stock market performance.

2.4 Conditional indexation schemes

As mentioned above, a defining property of conditional indexation schemes is that the guaranteed amount of pension is adjusted over time. For instance, if a participant is granted full indexation to wage inflation each year after her entry into the fund, then the two thresholds will converge with the guaranteed amount approaching the intention amount at retirement. On the other hand, if the participant is so unlucky as to receive no indexation at all during the entire working life, then the guaranteed amount at retirement will be same as the amount that was already guaranteed at the time of entry.

Given the diversity of conditional indexation schemes, it is far from being trivial to ask which one to take as the subject of welfare analysis. Our purpose in this chapter is to construct pension schemes with the defining property of dynamically adjusted guaranteed level as stylized conditional indexation schemes. The current practice of conditional indexation is implemented collectively by pension funds, and some studies show that the resulting conditional indexation schemes are not necessarily *financially fair* on a generation-by-generation basis in the sense that the value of the pension payoff may not be equal to the value of the contribution. It has been estimated that redistribution of wealth may reach 30% or more of the value of the liabilities [Kocken, 2007, p. 37] (cf. also Kocken [2006] for a more extensive theoretical analysis). For the purpose of welfare analysis, it is necessary to impose the condition of financial fairness, since the absence of financial fairness implies that one individual can be arbitrarily better off through wealth transfer from others. Therefore in the construction of stylized schemes, we impose the condition of financial fairness.

The construction of the stylized conditional indexation schemes starts with a digital contingent claim at time 0. Such a claim is optimal under a kinked utility function with $\gamma = 0$ and $\kappa = \infty$ (a piecewise linear function which drops to $-\infty$ at the guaranteed level and which saturates at the intention level). The claim is written on the stock index, and at retirement pays the guaranteed level θ_1, if the index is below a certain strike; otherwise it pays the intention level θ_2. At time 0, the value of the option is equal to the contribution value, W_0. In the Black-Scholes economy, the given θ_1, θ_2, and the option value W_0 determine the strike of the option as in (2.16). Browne [1999] shows that the policy to maximize the probability of reaching a given value wealth by a deadline is to buy a European digital option with a particular strike price and payoff. Applying Browne's insight, one can show that the digital claim maximizes the probability of reaching the intention level while subject to the constraint of not falling short of the guaranteed level. At time 0, the probability of reaching the intention level is determined by

$$p = \Phi\left[\lambda\sqrt{T} + \Phi^{-1}\left(\frac{W_0 e^{rT} - \theta_1}{\theta_2 - \theta_1}\right)\right] \tag{2.5}$$

where $\Phi^{-1}(\cdot)$ denotes the inverse of the standard normal cumulative distribution function. The above equation for p is well defined only for $e^{-rT}\theta_1 \leq W_0 \leq e^{-rT}\theta_2$, where the scheme is reduced to the all-bond scheme for $W_0 = e^{-rT}\theta_2$. In the following welfare analysis, the case where $W_0 > e^{-rT}\theta_2$ may arise. In such a case, we use the all-bond scheme with payoff $W_T = e^{-rT}W_0$ to replace both the digital scheme and the stylized conditional indexation scheme constructed on the basis of the digital scheme.

The digital option by itself reflects the idea of a guaranteed level, but not the idea of conditional indexation. We would like to increase the guaranteed level when circumstances allow. Assume that circumstances are indeed found favorable at a first review date following time 0; then it is possible to sell the digital option that was purchased at time 0 and to buy a new digital option that has an increased lower level. The self-financing property of the strategy is guaranteed by requiring that the value of the newly bought option at the time of its purchase is equal to the value of the previously owned option at that time.

In this way we obtain one constraint on the characteristics of the new option to be bought. However, a digital option is characterized by three parameters (upper level, lower level, and strike) so that two degrees of freedom remain. As a second constraint, we impose that the intention level θ_2 remains the same. The third constraint might be provided by imposing that the strike also remains the same, but a more basic requirement may be that the probability of reaching the upper level is kept constant. This may be motivated if one thinks of the size of the investment at time 0 as reflecting an implicit decision on the probability of reaching the intention level, via the relation (2.5).

Under the proposed rules, high returns on the stock market will result in an increase of the guaranteed level, while both the intention level and the probability of reaching that level remain constant. Under the assumptions of the Black-Scholes market, we can and will consider a continuous-time version of the proposed strategy; moreover, using the completeness of the Black-Scholes market, the process of buying and selling options can be replicated by suitable portfolio rebalancings. We will consider two versions of the proposed scheme; one in which the rules as stated above are applied irrespective of stock returns, so that also downward adjustments may take place, and another (which is closer to practice) in which adjustments in the downward direction are not made and one accepts that under adverse circumstances the probability of reaching the intention level decreases.

2.4.1 The updating rules

Version I: two-way adjustment This rule prescribes that the probability of reaching the intention level be constant over time. In particular, if the probability of reaching the intention level increases (decreases) due to an upturn (downturn) of the stock index, then the strike denoted by $K_t^{(1)}$ is adjusted upwards (downwards) to the level that restores the probability to the benchmark p. At the same time, the guaranteed level, denoted by $\theta_{1,t}^{(1)}$, is increased (decreased) to ensure that the update is self-financing. This scheme is mainly of academic interest; a more realistic scheme that allows only one-way adjustment is described below. Appendix 2.A shows that the two-way adjustment rule leads to dynamics of the strike $K_t^{(1)}$ and guaranteed level $\theta_{1,t}^{(1)}$ for $t < T$ given by

$$dK_t^{(1)} = \left[\frac{\Phi^{-1}(p)\sigma}{2\sqrt{T-t}} + \tfrac{1}{2}\sigma^2\right] K_t^{(1)}dt + \sigma K_t^{(1)}dZ_t, \quad K_0^{(1)} = K \tag{2.6}$$

$$d\theta_{1,t}^{(1)} = \frac{(\theta_2 - \theta_{1,t}^{(1)})\phi(d^{(1)}(t))}{[1-\Phi(d^{(1)}(t))]\sqrt{T-t}} \left(\tfrac{1}{2}\lambda dt + dZ_t\right), \quad \theta_{1,0}^{(1)} = \theta_1 \tag{2.7}$$

where

$$\begin{aligned} d^{(1)}(t) &= \frac{\log(S_t/K_t^{(1)}) + (r-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \\ &= \Phi^{-1}(p) - \lambda\sqrt{T-t}. \end{aligned}$$

As an alternative to the stochastic differential equation (2.6), the dynamics of the strike can be explicitly expressed as a function of the stock index value:

$$K_t^{(1)} = S_t \exp\left[(\mu - \tfrac{1}{2}\sigma^2)(T-t) - \Phi^{-1}(p)\sigma\sqrt{T-t}\right]. \tag{2.8}$$

As can be seen from (2.8), the strike tends to the stock index as t approaches T, which implies that the investment policy is to take an increasingly sensitive bet in the form of digital options that are increasingly *at the money*. For the dynamics of the guaranteed level, it can be shown that

$$\operatorname*{p-lim}_{t\uparrow T} \theta_{1,t}^{(1)} = \theta_2;$$

that is, the guaranteed levels converge in probability to the intention level as t approaches T (see Appendix 2.A). This is not surprising given that the drift and volatility terms of (2.7) "explode" as t approaches T unless the guaranteed level converges to the intention level.

Version II: ratchet adjustment This rule introduces a "ratchet" effect by allowing only upward adjustment of the guaranteed level. In a nutshell, the updating rule is to keep the probability of reaching the intention level no higher than the benchmark p. The updating is the same as version I in case of stock market upturns: if the stock markets rise, and the probability of reaching the intention level rises above the benchmark, then the strike and the guaranteed level are adjusted upwards in such a way as to restore the probability to the benchmark. In case of stock market downturns where the probability falls short of the benchmark, however, the strike and the guaranteed level are unchanged. As such, the strike, $K_t^{(2)}$, and the guaranteed level, $\theta_{1,t}^{(2)}$, can be adjusted upward only. The strike resulting from this version of conditional indexation, denoted by $K_t^{(2)}$, is the running maximum of the strike from version I (see Appendix 2.A), i. e.

$$K_t^{(2)} = \max_{s \leq t} K_s^{(1)}. \tag{2.9}$$

Given the dynamics of the strike, one can determine the updating of the guaranteed level by the requirement that the value of the updated pension rights should be unchanged. The SDE for the guaranteed level is not of a simple form, so it is omitted here. We only note that the running-maximum relationship does not hold for the guaranteed levels.

Figure 2.5 presents a simulated history of the strikes and guaranteed levels for the two updating rules. It illustrates the increasing volatility of the guaranteed level on the basis of the first version as t approaches T, and the running-maximum relationship between the strike prices resulting from the two rules.

2.4.2 The pension rights of the conditional indexation schemes

We now want to see the effect of both strategies on W_T, the realized capital at time T. First, define by continuity the values of the strikes and the guaranteed levels at time T for both updating rules, that is,

$$K_T^{(i)} \triangleq \lim_{t \uparrow T} K_t^{(i)}, \quad \theta_{1,T}^{(i)} \triangleq \lim_{t \uparrow T} \theta_t^{(i)}, \quad i = 1, 2 \tag{2.10}$$

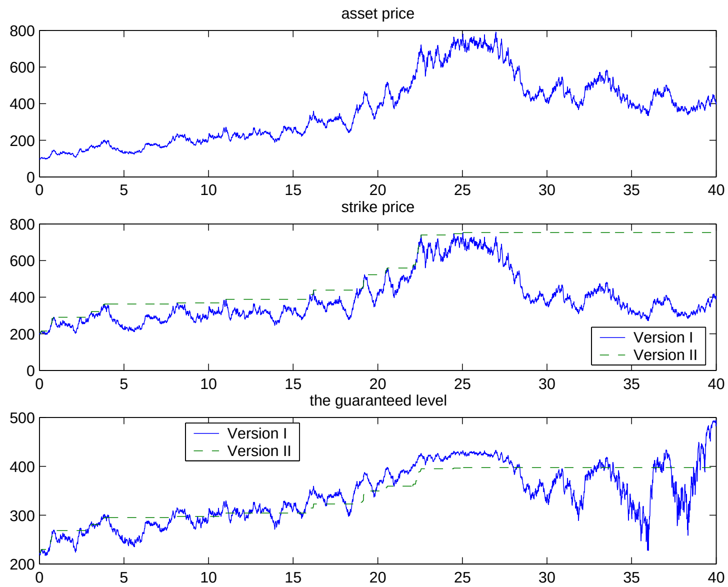

Figure 2.5: **Simulated histories of the strike and the guaranteed level** The upper panel is a simulated scenario of the stock index over 40 years, and the corresponding histories of the strike and the guaranteed level are in the middle and lower panels, respectively. In the middle and lower panels, the histories based on the two-way updating rule are represented by solid line, and those based on the ratchet updating by dashed line. For the parameter values of the simulation, see Table 2.1.

The pension rights resulting from the two versions of conditional indexation, $W_T^{(1)}$ and $W_T^{(2)}$, are then

$$W_T^{(i)} = (\theta_2 - \theta_{1,T}^{(i)}) 1_{S_T > K_T^{(i)}} + \theta_{1,T}^{(i)}, \quad i = 1, 2. \tag{2.11}$$

For the conditional indexation of version I, given that the guaranteed level converges to the intention level in probability, the pension rights at retirement also converge to the intention level. This is a peculiar outcome since it appears to construct an arbitrage opportunity; the proposed investment strategy seems to ensure a return that is higher than the riskless return. The explanation is that this version of conditional indexation allows "outrageous" investment behavior which violates the *admissibility* assumption in the finance literature (see, e. g. Section 6.C of Duffie [2001]). Like the well-known "doubling" strategy, the first version of conditional indexation involves shorting more and more of the riskless asset and going long in the risky asset in some states of nature, and it has to allow the possibility that pension asset value can go negative and be unbounded from below before the intention level is actually attained. Investment strategies of this nature are usually ruled out in the finance literature by the admissibility assumption which either prohibits the wealth process from going negative or imposes a square-integrability condition.

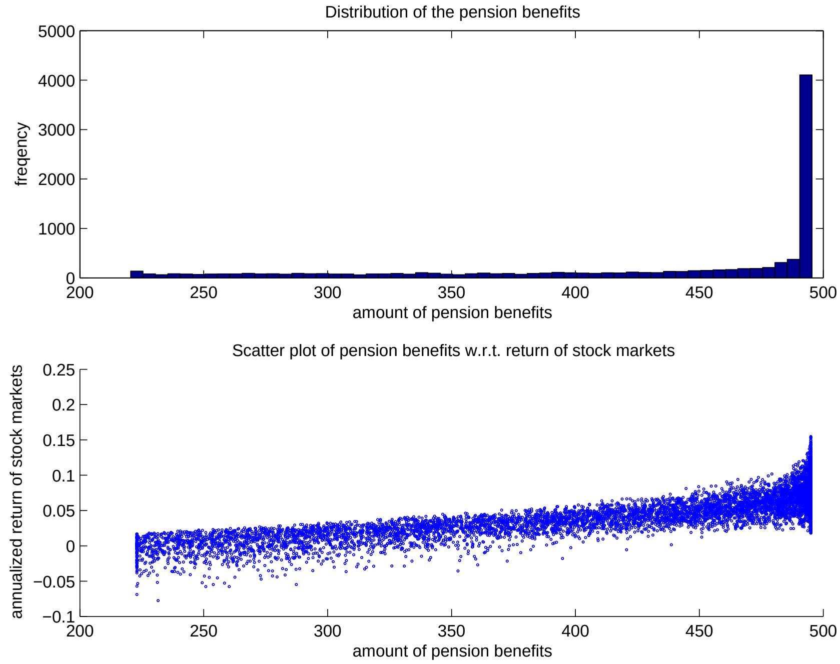

Figure 2.6: **The pension benefits resulting from the ratchet conditional indexation** The figure shows the pension profile from the ratchet rule of conditional indexation. The upper panel presents a simulated histogram of the pension rights whereas the lower panel illustrates the relation between the return on the stock index and the amount of pension rights. The number of simulations is 10,000. See Section 2.2.2 for the parameter values of the simulations.

For the ratchet indexation, Figure 2.6 presents the distribution of the pension rights in our standard example. Also shown is the relation between the pension rights and the annualized return on stock markets. In addition to the notable concentration on the two original thresholds, the pension rights generated by ratchet conditional indexation have considerable probability of falling in the intermediate region between the two thresholds. The pension rights are path-dependent, rather than dependent solely on the stock index level at retirement. The resultant pension scheme will be used as the stylized conditional indexation scheme in the following welfare analysis.

2.5 Welfare analysis

By now we have set up the criterion and the subject of the welfare analysis. To apply welfare analysis to the stylized conditional indexation scheme against the utility function (2.2), we have to resort to numerical methods because of the absence of an analytical solution. Therefore we need to choose the value of parameters characterizing the benchmark utility, in particular, κ and γ. Tversky and Kahneman [1992], based on psychological

Benchmark: $\gamma = 1$, $\kappa = \cdots$	1	2.25	5	10	100
optimal for $\gamma = 2$, $\kappa = 1$	22.2	3.7	8.7	21.9	86.1
best constant-proportion	0.0	3.4	4.2	5.6	9.2
corresponding RRA	1.0	1.5	3.0	3.5	5.5
optimal for $\gamma = 0$, $\kappa = \infty$	122.7	50.6	5.0	1.4	0.6
conditional indexation	122.7	29.8	9.9	6.4	6.0

Table 2.2: **The welfare loss of various schemes against the benchmark utility with different parameter values (in percent)** The table reports the welfare loss in terms of the additional percentage of contribution value which is required to for a pension scheme to reach the same level of expected utility as is obtained from the optimal strategy. The numbers in smaller font are the rates of relative risk aversion corresponding to the best constant-proportion schemes. Notice that the first column ($\kappa = 1$) actually uses the standard logarithmic utility as the benchmark.

experiments, estimate that the rates of local relative risk aversion are rather small (+0.12 in the gain region, and −0.12 in the loss region).[2] Most literature on application of the idea of loss aversion to finance (e. g. Ait-Sahalia and Brandt [2001], Barberis et al. [2001], Berkelaar et al. [2004], and Gomes [2005]) assumes that the coefficient of local relative risk aversion is between 0 and 1. We shall take $\gamma = 1$ in the following. As for κ, the estimated value of 2.25 that Tversky and Kahneman [1992] obtain in the context of loss aversion is on the basis of the choices of a group of individuals facing hypothetical decision problems. Given the critical importance of retirement income security, a higher degree of kinkedness may be reasonable. In addition, the standard CRRA utility will be included as a special case of (2.2). Thus we consider $\kappa \in \{1, 2.25, 5, 10, 100\}$.

The purpose of the welfare analysis is to assess the quality of conditional indexation schemes to individuals whose preference is characterized by the family of utility functions (2.2) in the context of lifecycle investment. The performance is measured by welfare loss in terms of the additional contribution value required for a pension scheme to reach the same level of expected utility as is obtained from the optimal strategy. Apart from the stylized conditional indexation scheme, other pension schemes are considered for comparison purposes (see Appendix 2.A for the computation of expected utility of these schemes).

[2] A utility function typical prospect theory is of the following convex-concave shape

$$U(W) = \begin{cases} -A(\theta - W)^{b_1} & \text{for } W \leq \theta, \\ +B(W - \theta)^{b_2} & \text{for } W > \theta. \end{cases}$$

It becomes approximately piecewise linear (i.e., approximately locally risk neutral) based on the estimated values of $b_1 = 0.88$ and $b_2 = -0.88$. As indicated by Sharpe [1998], the (approximate) local risk-neutrality leads to investment strategies that are rather extreme.

First consider the pension schemes resulting from constant-proportion strategies characterized by different stock weights. The second row of Table 2.2 shows results for a scheme with a constant stock weight of 50%, which, in the numerical setting as given in Table 2.1, is optimal for the CRRA preference when the coefficient of relative risk aversion is equal to 2. If the benchmark utility is the standard logarithmic utility ($\kappa = 1$), the welfare loss of this constant-proportion scheme is substantial because of insufficient risk taking: over 20% more contribution is needed for the scheme to obtain the same level of expected utility as the optimal strategy, which is a constant-proportion scheme with 100% stock weight. If the two reference points are present in the individual preference, and their significance is moderate with $\kappa = 2.25$, then this scheme looks much better with welfare loss less than 4%. The two kinks introduce first-order risk aversion, and hence make the 50%-stock-weight scheme look more favorable against the kinked benchmark utility than against the logarithmic benchmark. From this perspective, the parameters κ and γ, albeit reflecting different aspects of preference, substitute for each other to some extent. As the kinkedness parameter increases, however, the welfare loss increases considerably. A plausible explanation of large welfare loss for large κ value is that an individual with strongly kinked utility favors good downside protection at the cost of giving up upside potential, in which respect the constant-proportion scheme is poor.

In the sphere of constant-proportion strategies, one can vary the stock weight in order to maximize the expected utility with respect to the (possibly kinked) benchmark utility. We refer to the optimal constant-proportion schemes with respect to the benchmark utility as the *best constant-proportion* scheme. When the benchmark is reduced to the logarithmic utility, the best constant-proportion scheme has no welfare loss simply because the scheme keeps stock weight equal to level required by the logarithmic utility. As can be seen from the third row of Table 2.2, for kinked benchmark utility, the welfare loss of the best constant-proportion schemes is increasing with the kinkedness parameter. Welfare losses reflect the inability of constant-proportion schemes to provide downside protection required by the kinked benchmark utility. Moreover, for a more kinked benchmark, the best constant-proportion scheme decreases risk-taking, as is seen from the fact that the rate of relative risk aversion corresponding to the stock weight of the scheme (shown in smaller font in the table) is increasing with κ.

Another scheme we consider for comparison is the rigid digital scheme which at retirement pays the guaranteed level if the stock index is below the strike given by (2.16), and pays the intention level otherwise. As mentioned above, the digital scheme forms the basis of the construction of the stylized conditional indexation scheme, and is optimal with respect to the kinked utility function (2.2) with $\kappa = \infty$ and $\gamma = 0$. The digital scheme suffers from welfare loss in that (i) it assumes the greatest possible strength of the reference points ($\kappa = \infty$), and (ii) it assumes local risk neutrality ($\gamma = 0$). As shown in the fourth row of Table 2.2, the digital scheme's welfare loss decreases with the value of the kinkedness parameter κ used in the benchmark utility. When $\kappa \geq 10$, the welfare loss becomes rather small. It is a natural outcome, recalling that the pension profile which is optimal for a kinked benchmark utility with large values of κ resembles the payoff of a digital option as illustrated in Figure 2.3.

The stylized conditional indexation scheme, the focus of the welfare analysis, is built on the basis of the digital scheme through updating the lower threshold over time. As the bottom row of Table 2.2 shows, like the digital scheme, the conditional indexation scheme incurs a welfare loss which is decreasing with the kinkedness of the benchmark utility. For an individual with logarithmic or a mildly kinked utility ($\kappa = 1$, or 2.25), the utility loss of the conditional indexation scheme (and the digital scheme) is significant. Against the logarithmic benchmark, for the conditional indexation scheme and the digital scheme to achieve the same utility level as is obtained by the optimal scheme, it is insufficient that the contribution is increased to such a level that the probability of reaching the intention level is one, i. e. to the level equal to $\theta_2 e^{-rT}$. As mentioned earlier, the two schemes are assumed to be reduced to the all-bond scheme with payoff equal to $e^{-rT} W_0$, and the welfare loss is computed accordingly.

In comparison with the digital scheme, the dynamic updating of the guaranteed level introduces another element of suboptimality with respect to the benchmark utility, namely a non-constant lower reference point. It perhaps accounts for the welfare loss of the conditional indexation scheme being higher than that of the digital scheme in the case where $\kappa \geq 5$. Nevertheless, against the strongly kinked benchmark utility (e. g. $\kappa = 10$ or 100), the stylized conditional indexation scheme is close to the optimal as it has a moderate welfare loss of about 6%. In short, for individuals in whose preference the reference points play little role, the stylized conditional indexation scheme leads to material welfare loss, while for those paying much attention to the references, the scheme offers a reasonable option for retirement savings from the point of view of welfare analysis.

2.6 Concluding remarks

In this chapter we have used welfare analysis to evaluate the performance of conditional indexation schemes from a life-cycle investment perspective. This type of analysis is usually applied to study the effect of constraints. Conditional indexation in principle is not a constraint, but effectively such schemes imposed by compulsory participation cannot be undone by participants without costs, and they are taken as given by most people. For this reason, the welfare analysis on the basis of frictionless financial market, where participant actually need not care about the sub-optimality of pension schemes in that they are capable of undoing any financial contracts and achieving their optimal investment profile themselves free of cost, is still relevant in providing insights into the performance of conditional indexation schemes in practice.

For the purpose of this analysis, we needed to establish both the criterion and the subject of the utility analysis, namely a benchmark utility and representative conditional indexation schemes. To do justice to the two reference points which underlie the formulation of conditional indexation, and which often stand out in the discussion of pension provision, we propose as the benchmark utility an extended family of CRRA utility functions which accommodate the presence of reference points. The two reference points, as reflected by kinks in the utility function, lead to a pension investment policy with partial floor protection attained at the cost of forgoing some upside potential. The stylized conditional indexation scheme subject to welfare analysis is constructed to have a ratchet-adjusted guaranteed level, reflecting common practice. Possibly deviating from practice, the stylized scheme is financially fair because it is constructed on a self-financing basis. The property of financial fairness ensures that it makes sense to investigate conditional indexation schemes by means of utility analysis, and that such schemes are comparable to life-cycle investment policies.

Some numerical exercises show the influence of reference in the benchmark utility on the evaluation outcome of the stylized conditional indexation scheme. If the effect of the reference points is weak or even absent, the conditional indexation incurs substantial utility loss. If, however, the reference points are significant in preference, the scheme offers a reasonably good approach to pension provision as the welfare loss vis-à-vis the optimal is moderate. Conversely, increasing attention to conditional indexation may therefore be viewed as evidence of the presence and significance of reference points in participants' evaluation of pension provision.

The stylized conditional indexation scheme has been formulated in such a way that it can be implemented by an individual. However, the implementation of conditional indexation is usually done by collective funds. The collective implementation may play a role in saving transaction costs, and hence have an impact on the evaluation of pension schemes. It is a point beyond the scope of this chapter, and subject to further research. Another avenue for further research will be to use more realistic financial settings, for example, stochastic inflation rates, stochastic interest rates, and the possible long-term predictability of asset returns.

2.A Appendix

Derivation of the optimal pension profile W_T and investment strategy under piecewise power utility

The optimal pension profile W_T

The marginal utility function of the piecewise power utility (2.2) is[3]

$$U'(W) = \begin{cases} \kappa W^{-\gamma} & \text{for } W \leq \theta_1, \\ W^{-\gamma} & \text{for } \theta_1 \leq W \leq \theta_2, \\ \frac{1}{\kappa} W^{-\gamma} & \text{for } W \geq \theta_2. \end{cases}$$

Then the optimal pension profile at time T can be obtained by the Cox and Huang [1989] approach as expressed in (2.3);[4] in particular,

$$W_T = \begin{cases} (\kappa y \xi_T)^{-\frac{1}{\gamma}} & \text{for } \xi_T \leq \xi_1, \\ \theta_2 & \text{for } \xi_1 < \xi_T \leq \xi_2, \\ (y\xi_T)^{-\frac{1}{\gamma}} & \text{for } \xi_2 < \xi_T < \xi_3, \\ \theta_1 & \text{for } \xi_3 \leq \xi_T < \xi_4, \\ (\frac{y\xi_T}{\kappa})^{-\frac{1}{\gamma}} & \text{for } \xi_T \geq \xi_4, \end{cases} \tag{2.12}$$

where $\xi_1 = \frac{1}{\kappa y}\theta_2^{-\gamma}$, $\xi_2 = \frac{1}{y}\theta_2^{-\gamma}$, $\xi_3 = \frac{1}{y}\theta_1^{-\gamma}$ and $\xi_4 = \frac{\kappa}{y}\theta_1^{-\gamma}$.

[3]In the case of utility functions that are not everywhere differentiable, marginal utility can be expressed by the *superdifferential*, which is a multivalued function. For simplicity, we do not adapt the notation.

[4]Because the above marginal utility "function" involves a one-to-many mapping at θ_1 and θ_2, it is not a function in the normal sense of being one-to-one or many-to-one mapping, and is referred to as a *multivalued function*. The inverse relation (a generalization of the notion of inverse function) of this multivalued function is indeed a function in the normal sense, and the Cox and Huang [1989] approach still applies.

The optimal investment strategy

By substituting (2.12) into (2.4) and after some straightforward but somewhat tedious calculus, one can obtain the optimal wealth at time $0 \leq t < T$

$$\begin{aligned} W_t &= \theta_2 e^{-r(T-t)}[\Phi(d_1(\xi_2)) - \Phi(d_1(\xi_1))] + \theta_1 e^{-r(T-t)}[\Phi(d_1(\xi_4)) - \Phi(d_1(\xi_3))] \\ &\quad + (\kappa y \xi_t)^{-\frac{1}{\gamma}} e^{\Gamma(t)} \Phi(d_2(\xi_1)) + (\frac{y\xi_t}{\kappa})^{-\frac{1}{\gamma}} e^{\Gamma(t)} [1 - \Phi(d_2(\xi_4))] \\ &\quad + (y\xi_t)^{-\frac{1}{\gamma}} e^{\Gamma(t)} [\Phi(d_2(\xi_3)) - \Phi(d_2(\xi_2))], \end{aligned} \tag{2.13}$$

where $\Phi(\cdot)$ denotes the cumulative standard normal distribution function, and

$$\begin{aligned} d_1(x) &= \frac{\ln(x/\xi_t) + (r - \frac{1}{2}\lambda^2)(T-t)}{\lambda\sqrt{T-t}} \\ d_2(x) &= d_1(x) + \frac{\lambda\sqrt{T-t}}{\gamma} \\ \Gamma(t) &= \frac{1-\gamma}{\gamma}(r + \frac{\lambda^2}{2\gamma})(T-t). \end{aligned}$$

Applying the Itô rule to the expression of W_t (2.13), one can describe the optimal wealth process by a stochastic differential equation. Alternatively, one can in the Black-Scholes economy characterize any self-financing wealth process by the following stochastic differential equation

$$dW_t = (r + w_t\sigma\lambda)W_t dt + w_t \sigma W_t dZ_t,$$

where w_t denotes the stock weight at time t. Equating the diffusion parts of the two above-mentioned stochastic differential equations leads to the the optimal weight of risky assets at time $0 \leq t < T$

$$\begin{aligned} w_t^* &= \frac{\lambda}{\sigma W_t}\Bigg[\frac{\theta_2 e^{-r(T-t)}[\phi(d_1(\xi_2)) - \phi(d_1(\xi_1))] + \theta_1 e^{-r(T-t)}[\phi(d_1(\xi_4)) - \phi(d_1(\xi_3))]}{\lambda\sqrt{T-t}} \\ &\quad + (\kappa y\xi_t)^{-\frac{1}{\gamma}} e^{\Gamma(t)} \left(\frac{\Phi(d_2(\xi_1))}{\gamma} + \frac{\phi(d_2(\xi_1))}{\lambda\sqrt{T-t}}\right) \\ &\quad + (\frac{y\xi_t}{\kappa})^{-\frac{1}{\gamma}} e^{\Gamma(t)} \left(\frac{1 - \Phi(d_2(\xi_4))}{\gamma} - \frac{\phi(d_2(\xi_4))}{\lambda\sqrt{T-t}}\right) \\ &\quad + (y\xi_t)^{-\frac{1}{\gamma}} e^{\Gamma(t)} \left(\frac{\Phi(d_2(\xi_3)) - \Phi(d_2(\xi_2))}{\gamma} + \frac{\phi(d_2(\xi_3)) - \phi(d_2(\xi_2))}{\lambda\sqrt{T-t}}\right)\Bigg] \end{aligned}$$

where $\phi(\cdot)$ is the standard normal density function.

The optimal profile of pension rights in the special case where $\gamma = 0$ and $\kappa \to \infty$

For the extreme case where $\gamma = 0$ and $\kappa \to \infty$, the utility function (2.2) is reduced to

$$U(W) = \begin{cases} -\infty & \text{for } W < \theta_1, \\ aW + b & \text{for } \theta_1 \leq W \leq \theta_2, \\ a\theta_2 + b & \text{for } W \geq \theta_2, \end{cases} \tag{2.14}$$

As in the more general case, using the equivalent martingale approach of Cox and Huang [1989] for the utility can lead to optimal profile of pension rights contingent on the pricing kernel ξ_T:

$$W_T = \begin{cases} \theta_1 & \text{for } \xi_T \geq \xi_k, \\ \theta_2 & \text{for } \xi_T < \xi_k, \end{cases}$$

where

$$\xi_k = \exp\left[(\tfrac{1}{2}\sigma^2 - r)T + \lambda\sqrt{T}\Phi^{-1}\left(\frac{W_0 \exp(rT) - \theta_1}{\theta_2 - \theta_1}\right)\right],$$

where $\Phi^{-1}(\cdot)$ denotes the inverse of the standard normal cumulative distribution function. Given the one-to-one relation between the pricing kernel and the the stock index value the Black-Scholes model, we can express the optimal profile in terms of the stock index value S_T:

$$W_T = \begin{cases} \theta_1 & \text{for } S_T \leq K, \\ \theta_2 & \text{for } S_T > K, \end{cases} \tag{2.15}$$

where

$$K = S_0 \exp\left[(r - \tfrac{1}{2}\sigma^2)T - \sigma\sqrt{T}\Phi^{-1}\left(\frac{W_0 \exp(rT) - \theta_1}{\theta_2 - \theta_1}\right)\right]. \tag{2.16}$$

The dynamics of the strike and lower threshold in conditional indexation

Version I

Given the rule that the probability of reaching the intention level is unchanged over time, we have

$$\Pr\left(S_T > K_t^{(1)} | S_t, t\right) = p. \tag{2.17}$$

Conditioning on the stock index value S_t at time t, S_T can be expressed as

$$S_T = S_t \exp\left[(\mu - \frac{1}{2}\sigma^2)(T - t) + \sigma\sqrt{T - t}Z\right],$$

where Z is a standard normal variable. Inserting the expression into (2.17) generates the dynamics of the strike price as expressed by (2.8). Applying the Itô rule to (2.8), one obtains the characterization of the strike price by stochastic differential equation (2.6). Consider a European digital option which pays 0 at T if the stock price is lower than the strike, and pays 1 otherwise. We allow the strike K_t to change over time, and denote the pricing formula of the digital option by $F(K_t, S_t, t)$. Given the dynamics of the strike price from version I of conditional indexation, the pricing formula of the digital option $F(K_t^{(1)}, S_t, t)$ is simply a deterministic function $F_1(t)$

$$F(K_t^{(1)}, S_t, t) = F_1(t) = e^{-r(T-t)}\Phi\left(d^{(1)}(t)\right),$$

where

$$d^{(1)}(t) = \Phi^{-1}(p) - \lambda\sqrt{T-t}.$$

For the dynamics of the guaranteed level, consider the updating in discrete time first. At time t, the pension value is

$$W_t = (\theta_2 - \theta_{1,t}^{(1)})F(K_t^{(1)}, S_t, t) + \theta_{1,t}^{(1)}e^{-r(T-t)}.$$

At the time $t + \Delta t$ *before* applying conditional indexation, the pension value changes to

$$W_{t+\Delta t} = (\theta_2 - \theta_{1,t}^{(1)})F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t) + \theta_{1,t}^{(1)}e^{-r(T-t-\Delta t)}.$$

The pension value *after* conditional indexation is

$$W'_{t+\Delta t} = (\theta_2 - \theta_{1,t+\Delta t}^{(1)})F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t) + \theta_{1,t+\Delta t}^{(1)}e^{-r(T-t-\Delta t)}.$$

Since the conditional indexation does not change the pension value at time $t + \Delta t$, we have $W_{t+\Delta t} = W'_{t+\Delta t}$ and then the adjustment of $\theta_{1,t}^{(1)}$ is given by

$$\Delta\theta_{1,t}^{(1)} = \theta_{1,t+\Delta t}^{(1)} - \theta_{1,t}^{(1)} = \frac{(\theta_{1,t}^{(1)} - \theta_2)[F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t)]}{e^{-r(T-t-\Delta t)} - F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t)}.$$

One can decompose the term $F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t)$ in the numerator as

$$\begin{aligned} F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t) &- F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t) \\ &= \left[F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_t, t)\right] \\ &\quad - \left[F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_t, t)\right]. \end{aligned} \tag{2.18}$$

For the first term on the right-hand side of (2.18), we have

$$\begin{aligned} F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_t, t) &= F_1(t+\Delta t) - F_1(t) = \frac{dF_1}{dt}(t)\Delta t + o(\Delta t) \\ &= \left[e^{-r(T-t)} \frac{\lambda}{2\sqrt{T-t}} \phi\left(d^{(1)}(t)\right) + re^{-r(T-t)}\Phi(d^{(1)}(t)) \right] \Delta t + o(\Delta t), \quad (2.19) \end{aligned}$$

where $o(\cdot)$ denotes the higher order term. For the second on the right-hand side of (2.18), $F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_t, t)$, the difference is taken on the basis of fixed strike price, and hence is the difference of the value of a digital option with fixed strike price. We use $F_2(S_t, t)$ to denote the digital option pricing formula with a fixed strike K, and

$$F_2(S_t, t) = e^{-r(T-t)}\Phi(d(t)),$$

where

$$d(K, t) = \frac{\log(S_t/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

On the basis of the Itô rule, one has

$$\begin{aligned} dF_2(S_t, t) &= \frac{\partial F_2}{\partial t}(S_t, t)dt + \frac{\partial F_2}{\partial S_t}(S_t, t) + \frac{1}{2}\frac{\partial^2 F_2}{\partial S_t^2}(S_t, t)d[S, S]_t \\ &= \left[e^{-r(T-t)} r\Phi(d(t)) + \lambda \frac{e^{-r(T-t)}\phi(d(t))}{\sqrt{T-t}} \right] dt + \frac{e^{-r(T-t)}\phi(d(t))}{\sqrt{T-t}} dZ_t. \end{aligned}$$

Because $dF_2(S_t, t)$ is equal to $F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_t, t)$, and because for $K = K_t^{(1)}$, $d(K, t) = d^{(1)}(t)$, one can write

$$\begin{aligned} &F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_t, t) \\ &= \left[e^{-r(T-t)} r\Phi(d^{(1)}(t)) + \lambda \frac{e^{-r(T-t)}\phi(d(t))}{\sqrt{T-t}} \right] \Delta t + \frac{e^{-r(T-t)}\phi(d^{(1)}(t))}{\sqrt{T-t}} \Delta Z_t. \quad (2.20) \end{aligned}$$

Substituting (2.19) and (2.20) into (2.18), one gets

$$F(K_{t+\Delta t}^{(1)}, S_{t+\Delta t}, t+\Delta t) - F(K_t^{(1)}, S_{t+\Delta t}, t+\Delta t) = \frac{(\theta_2 - \theta_{1,t})\phi(d^{(1)}(t))}{[1 - \Phi(d^{(1)}(t))]\sqrt{T-t}} \left(\frac{\lambda}{2}\Delta t + \Delta Z_t \right),$$

which in continuous time converges to (2.7). Thus at last we obtain the stochastic differential equation characterizing the evolution of the guaranteed level in continuous time.

Next we shall show that the guaranteed level reaches the intention level before time T almost surely. Consider first the stochastic process

$$dX_t = -\frac{X_t}{\sqrt{T-t}} \left(\tfrac{1}{2}\lambda dt + dZ_t \right), \quad X_0 = \theta_2 - \theta_{1,0}.$$

Applying the Itô rule to the logarithmic transformation $Y_t = \log X_t$ leads to

$$dY_t = -\frac{1}{2}\left(\frac{\lambda}{\sqrt{T-t}} + \frac{1}{T-t}\right) dt - \frac{1}{\sqrt{T-t}} dZ_t,$$

or

$$Y_t = Y_0 + \lambda(\sqrt{T-t} - \sqrt{T}) + \tfrac{1}{2}\log(T-t) - \tfrac{1}{2}\log T - \int_0^t \frac{1}{\sqrt{T-u}} dZ_u.$$

Thus one has

$$X_t = X_0 \exp\left[\lambda(\sqrt{T-t} - \sqrt{T})\right] \frac{\sqrt{T-t}}{\sqrt{T}} \exp\left(-\int_0^t \frac{1}{\sqrt{T-u}} dZ_u\right).$$

Karatzas and Shreve [1998] (Section 1.2) show that under the time change $t = T - Te^{-s}$, $\int_0^t \frac{1}{\sqrt{T-s}} dZ_s$ is a Brownian motion defined for $0 \le s < \infty$. The time-changed stochastic process is

$$\hat{X}_s = X_0 \exp\left[\lambda(\sqrt{Te^{-s}} - \sqrt{T})\right] \exp\left(Z_s - \frac{s}{2}\right).$$

Consequently,

$$\operatorname*{p-lim}_{t\uparrow T} X_t = \operatorname*{p-lim}_{s\uparrow\infty} \hat{X}_s = 0,$$

where the notation “p-lim” denotes convergence in probability. Thus, for the stochastic process $\theta_{1,t}$ defined by

$$d\theta_{1,t} = \frac{\theta_2 - \theta_{1,t}}{\sqrt{T-t}} \left(\tfrac{1}{2}\lambda dt + dZ_t\right), \quad \theta_{1,0} = \theta_1 \tag{2.21}$$

it holds that $\text{p-lim}_{t\uparrow T}\, \theta_{1,t} = \theta_2$.

We now turn to the case in which $\theta_{1,t}$ is defined by (2.7). Multiplying the right hand side of (2.21) by $\frac{\phi(d^{(1)}(t))}{1-\Phi(d^{(1)}(t))}$ will result in (2.7). Given that the term $\frac{\phi(d^{(1)}(t))}{1-\Phi(d^{(1)}(t))}$ is continuous, and tends to a finite constant as time approaches T, it also holds that $\text{p-lim}_{t\uparrow T}\, \theta_{1,t} = \theta_2$ in this case.

Version II

In this appendix, we shall motivate the rule of conditional indexation as given by (2.9) through the case of discrete updating. First of all, by definition one has

$$K_0^{(2)} = \max(K_0^{(1)}, K_0^{(1)}).$$

Next we show that the running-maximum relationship holds for any time $t + \Delta t$ through mathematical induction. Assume that the statement holds for time t, i. e.

$$K_t^{(2)} = \max_{s\le t} K_s^{(1)}.$$

The probability of reaching the intention level at time $t + \Delta t$ before updating the strike is

$$p_{t+\Delta t} = \Phi\left[(\mu - \tfrac{1}{2}\sigma^2)(T - t - \Delta t) - \frac{\ln(K_t^{(2)}/S_{t+\Delta t})}{\sigma\sqrt{T - t - \Delta t}}\right],$$

and the strike price is updated by

$$K_{t+\Delta t}^{(2)} = S_{t+\Delta t}\exp\left[(\mu - \tfrac{1}{2}\sigma^2)(T - t - \Delta t) - \Phi^{-1}[\min(p, p_{t+\Delta t})]\sigma\sqrt{T - t - \Delta t}\right].$$

Given this assumption, one can verify that

$$\begin{aligned}
K_{\Delta t}^{(2)} &= \max\left\{K_{t+\Delta t}^{(1)}, S_{t+\Delta t}\exp\left[(\mu - \tfrac{1}{2}\sigma^2)(T - t - \Delta t) - \Phi^{-1}(p_{t+\Delta t})\sigma\sqrt{T - t - \Delta t}\right]\right\} \\
&= \max\left\{K_{t+\Delta t}^{(1)}, K_t^{(2)}\right\} \\
&= \max\left\{K_{t+\Delta t}^{(1)}, \max_{0\le s\le t} K_s^{(1)}\right\} \\
&= \max_{0\le s\le t+\Delta t} K_s^{(1)}.
\end{aligned}$$

Thus we have proven that for discrete updating, the strike price from version II is the running maximum of that from version I, which motivates the rule given by (2.9) in continuous time.

Expected utility computation

The appendix shows the computation of the expected utility of the schemes studied in Section 2.5 with respect to the piecewise power utility. For the scheme with the ratchet conditional indexation, the expected utility can be obtained through Monte Carlo simulations. For the other schemes, analytical formulae can be obtained through integral calculus.

The expected utility of the optimal strategy

$$\begin{aligned}
EU(W_T) &= -\frac{1}{1-\gamma} + \frac{\theta_1^{1-\gamma}}{1-\gamma}[1 - \Phi(c_1(\xi_3))] + \frac{\theta_2^{1-\gamma}}{1-\gamma}\Phi(c_1(\xi_2)) \\
&\quad - \frac{\kappa\theta_1^{1-\gamma}}{1-\gamma}[1 - \Phi(c_1(\xi_4))] - \frac{\theta_2^{1-\gamma}}{\kappa(1-\gamma)}\Phi(c_1(\xi_1)) \\
&\quad + \frac{1}{\kappa(1-\gamma)}(\kappa y)^{\frac{\gamma-1}{\gamma}} e^{\Gamma(0)}\Phi(c_2(\xi_1)) + \frac{\kappa}{1-\gamma}(\frac{y}{\kappa})^{\frac{\gamma-1}{\gamma}} e^{\Gamma(0)}[1 - \Phi(c_2(\xi_4))] \\
&\quad + \frac{1}{1-\gamma} y^{\frac{\gamma-1}{\gamma}} e^{\Gamma(0)}[\Phi(c_2(\xi_3)) - \Phi(c_2(\xi_2))],
\end{aligned}$$

where

$$
\begin{aligned}
c_1(x) &= \frac{\ln(x) + (r + \frac{1}{2}\lambda^2)T}{\lambda\sqrt{T}} \\
c_2(x) &= \frac{\ln(x) + (r + \frac{1}{2}\frac{1+\gamma}{1-\gamma}\lambda^2)T}{\lambda\sqrt{T}} \\
\Gamma(0) &= \left(r + \frac{1}{1-\gamma}\frac{\lambda^2}{2}\right)\frac{\gamma}{1-\gamma}T
\end{aligned}
$$

The expected utility of the digital scheme which is optimal for the utility with $\kappa = \infty$ and $\gamma = 0$

As is shown in Appendix 2.A, if the benchmark utility function (2.2) has the following parameter values: $\kappa = \infty$ and $\gamma = 0$, then the optimal pension profile reduces to a digital option. Against the general form of utility function (2.2), the expected utility of the digital scheme is

$$EU(W_T) = \frac{\theta_1^{1-\gamma} - 1}{1-\gamma}(1-p) + \frac{\theta_2^{1-\gamma} - 1}{1-\gamma}p,$$

where p is determined by (2.5).

The expected utility of the constant-proportion strategy

Let γ_c denote the coefficient of relative risk aversion underlying a constant-proportion strategy. The stock weight of the constant-proportion strategy is

$$w^c = \frac{\mu - r}{\gamma_c \sigma^2},$$

and the pension rights are

$$W_T^c = W_0 \exp\left[[w^c\mu + (1-w^c)r - \frac{1}{2}(w^c\sigma)^2]T + w^c\sigma Z_T\right].$$

The expected utility derived from the constant-proportion strategy with stock weight w^c is

$$
\begin{aligned}
EU(W_T^c) = & -\frac{1}{1-\gamma} + (1-\kappa)\frac{\theta_1^{1-\gamma}}{1-\gamma}[1 - \Phi(c_1(\xi_1^c))] + (1 - \frac{1}{\kappa})\frac{\theta_2^{1-\gamma}}{1-\gamma}\Phi(c_1(\xi_2^c)) \\
& + \frac{\kappa}{1-\gamma}W_0^{1-\gamma}\exp[H - (1-\gamma)\Gamma_c(0)][1 - \Phi(c_3(\xi_1^c))] \\
& + \frac{1}{\kappa(1-\gamma)}W_0^{1-\gamma}\exp[H - (1-\gamma)\Gamma_c(0)]\Phi(c_3(\xi_2^c)) \\
& + \frac{1}{1-\gamma}W_0^{1-\gamma}\exp[H - (1-\gamma)\Gamma_c(0)][\Phi(c_3(\xi_1^c)) - \Phi(c_3(\xi_2^c))],
\end{aligned}
$$

where

$$
\begin{aligned}
\Gamma_c(0) &= \left(r + \frac{1}{1-\gamma_c}\frac{\lambda^2}{2}\right)\frac{\gamma_c}{1-\gamma_c}T \\
\xi_1^c &= \left(\frac{W_0}{\theta_1}\right)^{\gamma_c}\exp\left[-\gamma_c\Gamma_c(0)\right] \\
\xi_2^c &= \left(\frac{W_0}{\theta_2}\right)^{\gamma_c}\exp\left[-\gamma_c\Gamma_c(0)\right] \\
H &= \frac{1-\gamma}{\gamma_c}\left[r + \frac{1+\gamma_c-\gamma}{2\gamma_c}\lambda^2\right]T \\
c_3(x) &= \frac{\ln(x) + (r + \frac{1}{2}\lambda^2 + \frac{1-\gamma}{\gamma_c}\lambda^2)T}{\lambda\sqrt{T}}
\end{aligned}
$$

Chapter 3

Valuation of Contingent Pension Liabilities and Implementation of Conditional Indexation

3.1 Introduction

In the insurance and pension industries, some contracts are formulated in such a way that the payoff to beneficiaries is linked to the financial status of insurers or pension funds. In insurance, a typical example is given by participating life insurance policies (also known as with-profits policies), which provide a guaranteed return and allow beneficiaries to participate in the profit of a life insurance company through a profit-sharing scheme. Participating contracts make up a significant part of the life insurance market in many industrial countries. In the field of pensions, there exists a similar practice known as *conditional indexation*, which links the pension benefits to the financial position of a pension fund. The practice of conditional indexation has been adopted by many pension funds in the Netherlands in recent years, and its introduction is under discussion in the UK.

As already stated in the Introduction, a conditional indexation scheme typically grant inflation compensation to its participants on the basis of its *funding ratio*, the ratio of pension asset value to pension liability value. In particular, the level of compensation for inflation (indexation) applied in a certain year is determined according to a rule of the following form: (i) if the funding ratio is below some threshold (e. g. 110%), there is no indexation to inflation; (ii) if the funding ratio is above some upper threshold (e. g. 140%), the pension rights will be fully indexed to inflation; (iii) if the ratio is in between, some intermediate level of indexation will apply. The rule that links the funding ratio to the indexation decision is know as a "policy ladder". In practice, such rules are also subject to the discretion of pension board.

Against the backdrop of the international move promoted by the International Accounting Standards Board (IASB) towards the market-based, fair value accountancy standard, the valuation of these contracts contingent on the financial position of insurers and pension funds have become a subject of increasing interest both to academics and practitioners. For participating life insurance contracts, the valuation has been analyzed in the contingent claim pricing framework by a number of recent papers, for example, Aase and Persson [1997], Grosen and Jørgensen [2002], Bauer, Kiesel, Kling, and Ruß [2006], Ballotta, Haberman, and Wang [2006], Gatzert and Kling [2007], and Kleinow and Willder [2007]. In this chapter, we consider the valuation of conditionally indexed pension liabilities in the same framework. We also note that a good understanding of how to value this type of pension liabilities is also needed in dealing with pension rights transfer of conditional indexation schemes.

The conditional indexation of pension benefit may lead to some interesting and challenging issues in the market valuation of pension liabilities. First of all, the value of liabilities and the funding ratio have to be determined *simultaneously* for conditional indexation schemes. The reason is that the indexation level for the current year depends on the funding ratio via a policy ladder, and at the same time, the indexation level has a feedback effect on the liability value and in turn on the funding ratio through market valuation. In other words, a *circularity* problem arises in the valuation of conditionally indexed pension liabilities:

$$\boxed{\textbf{Funding ratio} = \tfrac{\text{Asset}}{\text{Liability}}} \;\underset{\textit{conditional indexation}}{\overset{\textit{market valuation}}{\longleftrightarrow}}\; \boxed{\textbf{Indexation level}}$$

Therefore, the determinations of the indexation level, liability value, and funding ratio are interdependent, and hence need to be addressed simultaneously through looking for a coherent solution in the sense that the indexation level and the funding ratio are consistent with each other with respect to the policy ladder and the market valuation approach.

Another complication arises from the *intertemporal* dimension of market valuation of pension liabilities. The lion's share of pension liabilities is from pension rights payable in the future. Under conditional indexation, the future pension payments can be thought of as collar-structured contingent payoffs: the payoffs are subject to floors and caps determined by minimum and maximum indexation, and contingent on the future funding ratio. In principle, one can reach a market valuation of the liabilities using contingent payoff pricing techniques developed along the lines of Black and Scholes [1973] and Merton [1973]. A complication resulting from conditional indexation is that *pension payments to be made at different time points are inter-dependent*. To see the intertemporal dependence, consider a pension fund that is to make pension payment at only two future dates (i. e. there are two European options with different expiry dates). The intertemporal dependence may arise in two ways. Firstly, the first pension payment, which can be said to be the actual payoff of the first option in the language of financial derivatives, will reduce the pension asset value. As a consequence, it has an impact on the future funding ratio, in terms of which the second payment (or the second option) is defined. Therefore the actual payoff of the first option has an impact on the *underlying* of the second, and in turn an impact on the valuation of the second. Due to the intertemporal dependence, these two options need to be considered together for their valuation, rather than one by one as the usual case of pricing a portfolio of options where the underlying dynamics is (or is assumed to be) not affected by the exercise of component options.

The second source of the intertemporal dependence is the practice that indexation is accumulated over time. For example, if participants are granted full indexation in 2005 and retirees receive pension benefits increased by the inflation, then the minimum amount of pension in 2006 is the amount paid in the previous year whatever the indexation level is decided for 2006 (assuming positive inflation). In the language of options, the year-by-year accumulation of conditional indexation leads to a case where the actual payoff of earlier options affects the *parameters* characterizing later options (the lower bound of pension rights in this context), i.e. payoff depends on state variables other the current funding ratio. In other words, the parameters which define the options are specified by some function (mechanism), but not specified as a known number as in most cases in options pricing.

When the indexation is accumulated year by year, referred to as the *cumulative* case in the following, there exists intertemporal dependence through both the underlying and the parameters. In comparison with the intertemporal dependence through the underlying, the dependence through the parameters results in stronger path-dependence in option pricing, and hence leads to a new dimension of complication in the value of pension liabilities as will be shown below. To distinguish these two types of intertemporal dependence, we discuss a case in which indexation is *not* accumulated and hence there only exists the intertemporal dependence through the underlying. And this case is referred to as the *noncumulative* case in this chapter.

Note that the circularity problem arises from our assumption that the funding ratio used for indexation decision is computed using the true liability value. By this assumption, we actually impose the requirement that the funding ratio used for indexation decisions is also reliable as an indicator of the financial soundness of pension funds. For this reason, we refer to as the *consistent* implementation of conditional indexation the case where the funding ratio used in indexation decisions is based on the true value of liabilities.

A way to avoid solving circularity problems is to replace the true value of liabilities (including the value of the indexation options) by a value that is easier to compute, such as the value of the liabilities without indexation, and to assume that indexation will be based on the funding ratio as computed from that value. Current practice at pension funds implementing conditional indexation is actually of this type. Similar assumptions are made in the recent papers Nijman and Koijen [2006] and de Jong [2008], who address the valuation of the liabilities of conditional indexation schemes. Both papers focus on the inflation risk stemming from the link to inflation which is uncertain, while abstracting from the circularity problem and intertemporal dependence inherent in the consistent implementation of conditional indexation. Similar to the current practice of conditional indexation, profit-sharing decisions in participating life insurance contracts are based on the *book* value, rather than the market value, of such contracts, and hence there does not exist the circularity problem either. So in the analysis of participating contracts as carried out in the papers mentioned earlier, no circularity problems need to be solved.

The procedure in which a proxy of funding ratio, rather than that based on the true value of liabilities, is used for indexation decisions is referred to as the *proxy-based* implementation of conditional indexation in this chapter. Though circumventing the circularity problem, this type of implementation leads to another complication. The funding ratio proxy assuming a fixed indexation level may not reflect the financial soundness of pension funds from the perspective of contingent claims pricing, since in fact the indexation levels to be decided in the future are contingent and varied. Therefore the implementation using a proxy calls for two funding ratios in the system: a proxy for indexation decisions, and the *actual* funding ratio for measuring the financial soundness. In the absence of the circularity problem, the actual values can be obtained by applying classical option pricing techniques. As will be shown below, for a given funding ratio proxy, the actual funding differs because the investment policies of pension funds differ, because the policy ladders differ, and because the demographic composition differs. It may be a point of concern from the regulatory perspective, if only the funding ratio proxy assuming a fixed indexation level is reported.

Presumably, the very fundamental idea underlying conditional indexation is that a pension fund pays more to its participants when it possesses more *vis-à-vis its liability*. Because a fixed-indexation proxy presumes indexation to be constant, but the actual indexation will be contingent and varied, a fixed-indexation proxy is hardly a right measure of how much a fund possesses vis-à-vis its liability. Therefore the proxy-based implementation in principle cannot ensure that this fundamental idea is actually enforced, and it allows the scenario that a pension fund pays out more even when it has less relative to its liability. Given the inconsistency between the intended idea and the actual implementation, one might ask what consequences the proxy-based implementation may have.

We proceed as follows. Section 3.2 presents the model and the computational procedure for the consistent implementation. To illustrate the computational procedure and show the impact of investment strategies and policy ladders on the liability value, we present some numerical examples in Section 3.3. Assuming the proxy-based implementation, Section 3.4 shows how the discrepancy between the funding ratio proxy and the actual funding ratio depends on investment policies and policy ladders. Some concluding observations are in Section 5.6.

3.2 Consistent implementation and computational procedure

3.2.1 The economy and the pension fund

Although the computation procedure we present below applies in more general settings, we consider for simplicity the standard Black-Scholes economy. In particular,

- the only risk factor is stock market risk, and it is traded through a stock index S_t following geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dZ_t,$$

where μ and σ are the constant drift and volatility parameters, and Z_t is the standard Brownian motion.

- The riskless asset is a cash bond with constant interest rate r, whose price, B_t, changes according to

$$dB_t = rB_t dt.$$

- The inflation rate, ρ, is constant.

We consider a pension fund which is liable for making pension payments at N time points: $T_1, \ldots, T_N$, and which will be liquidated at the terminal time T_N. In the noncumulative case, the payment at T_i $(i = 1, \ldots, N)$ is subject to a decision to be made at T_i on the indexation level applied over the period from the initial time T_0 to the current time T_i. In the cumulative case, however, the indexation decision at time T_i is made with respect to the inflation over the immediately previous period from T_{i-1} to T_i. Thus the payment at T_i is determined by the current indexation at time T_i, as well as by all previous indexation decisions made times from T_1 through T_{i-1}. We are now at time t $(< T_1)$ and to compute the liability value and funding ratio. Note that the model is based on projected benefit obligations rather than accrued benefit obligations.

The indexation decision is made according to a policy ladder, a formula that produces a indexation level applied over a certain period to the pension rights paid within a short time (usually one year in practice) on the basis of the current funding ratio.[1] Given the short interval between the time when the indexation is determined on the basis of the current funding ration, and the time when the pension rights with the decided indexation level are actually paid out, it is reasonable to model the policy ladder as a formula that determines the pension rights paid *immediately*, denoted by L^c, on the basis of the current funding ratio.

As mentioned in the introduction, the policy ladders in practice distinguish three regions: minimum, partial, and maximum indexation. For concreteness, we consider a piecewise linear form:

$$L^c = \begin{cases} L_\ell & \text{for } FR < K_\ell, \\ L_\ell + \beta(FR - K_\ell) & \text{for } K_\ell \leq FR \leq K_u, \\ L_u & \text{for } FR > K_u, \end{cases} \tag{3.1}$$

[1] In an average-salary scheme, the indexation level also apply to active workers, but we abstract from the added complication from the practice.

where FR is the current funding ratio, K_u and K_ℓ are the upper and lower thresholds of funding ratio, L_u and L_ℓ are the pension rights with maximum and minimum indexation levels, and $\beta = \frac{L_u - L_\ell}{K_u - K_\ell}$ to ensure that the function is continuous. For fixed L_u, L_ℓ and K_ℓ, the value of K_u (or β) determines how generous the policy ladder is. Note that in the noncumulative case, L_u and L_ℓ are known parameters while in the cumulative case, they are variables to be determined.

The pension fund is assumed to adopt a constant-proportion investment strategy. That is, the pension fund keeps the asset value invested in stocks (represented here by the stock index) a constant fraction of the total asset value, and we denote the constant proportion by α.

3.2.2 Computational procedure

The circularity problem

We consider the consistent implementation, where the funding ratio used in indexation conditions is computed based on the true value of the liabilities. For the computation of the funding ratio, we can decompose the pension liability value at any time t into two components: the value of pension right paid immediately L_t^c, and the value of pension rights to be paid in the future, denoted by L_t^f. We can write

$$FR = \frac{A}{L^c + L^f}, \tag{3.2}$$

where A is the current asset value, and the time subscript t is dropped for simplicity. The distinction between L^c and L^f relies on the time when the liability value and funding ratio are computed. For instance, L^c is zero and all the liability value is captured by L^f at time t, whereas L^c is the actual payment and L^f is zero at the terminal time T_N.

Equations (3.1) and (3.2) characterize the circularity problem in the consistent implementation of conditional indexation. From these two equations it follows that:

$$L^c = \begin{cases} L_\ell & \text{for } A < (L_\ell + L_f)K_\ell, \\ \frac{1}{2}\Big[L_\ell - L_f - \beta K_\ell + & \\ \quad \sqrt{4\beta A + (L_f + L_\ell - \beta K_\ell)^2}\Big] & \text{for } (L_\ell + L_f)K_\ell \le A \le (L_u + L_f)K_u, \\ L_u & \text{for } A > (L_u + L_f)K_u, \end{cases}$$

or in more compact form,

$$L^c = \min\left\{\max\left[\frac{1}{2}\left(L_\ell - L_f - \beta K_\ell + \sqrt{4\beta A + (L_f + L_\ell - \beta K_\ell)^2}\right), L_\ell\right], L_u\right\}. \quad (3.3)$$

The above equation solves the circularity problem by finding the coherent solution in the sense that the pension rights paid immediately and the current funding are consistent.

We shall describe the procedure to compute the liability value at t, and then through some numerical example to show how the liability value depends on the asset, and the impact of investment strategies and policy ladders on the liability value and funding ratio. First, let us introduce the following notations we need:

- A_τ: the asset value at time τ *before* pension payment
- A'_τ: the asset value at time τ *after* pension payment
- L_τ: the liability value of the pension fund at time τ *before* pension payment if there is any
- $\mathcal{L}_\tau(\cdot)$: the function associating the state variable(s) to L_τ
- L^ℓ_τ: the pension paid at τ with *minimum* indexation
- L^u_τ: the pension paid at τ with *maximum* indexation
- L^c_τ: the pension made immediately from the vantage point at time τ
- $\mathcal{L}^c_\tau(\cdot)$: the function relating the state variable(s) to L^c_τ
- L^f_τ: the pension to be made in the future from the vantage point at time τ
- $\mathcal{L}^f_\tau(\cdot)$: the function relating the state variable(s) to L^f_τ
- G: unconditional nominal benefit

The liability value at time t is computed using *backward* method. We note that the computation procedure is similar to that for pricing Bermudan options. The main difference is that for Bermudan options, we solve an optimization problem at each early exercise date, whereas for the liability valuation, we solve the circularity problem at each pension payment date. We shall distinguish two types of intertemporal dependence. For ease of exposition, we start with the noncumulative case, and then address the cumulative case.

The noncumulative case

Observe first that in the noncumulative case, the pensions paid at time T_i ($i = 1, \ldots, N$) with minimum and maximum indexation are given by

$$L^\ell_{T_i} = G, \quad L^u_{T_i} = G\exp[\rho(T_i - T_0)].$$

The recursive approach starts from the terminal time T_N, when the pension fund makes the last payment, $L^c_{T_N}$, and then it is liquidated. It is assumed that at the terminal time, the deficit short of the guaranteed level is made up for a third party, for instance, the pension regulator, whereas the surplus beyond the full indexation pension payoff will go to the third party for free. Given $L^f_{T_N} = 0$, we can obtain $L^c_{T_N}$ by applying (3.3) to solve the circularity problem at time T_N:

$$L^c_{T_N} = \min\left\{\max\left[\frac{1}{2}\left(L^\ell_{T_N} - \beta_{T_N} K_\ell + \sqrt{4\beta A_{T_N} + (L^\ell_{T_N} - \beta K_\ell)^2}\right), L^\ell_{T_N}\right], L^u_{T_N}\right\}. \tag{3.4}$$

where $\beta_{T_N} = \frac{L^u_{T_N} - L^\ell_{T_N}}{K_u - K_\ell}$. As can be seen from the above equation, the last payment is dependent on the asset value at T_N, A_{T_N}, and hence we can write

$$L^c_{T_N} = \mathcal{L}^c_{T_N}(A_{T_N}). \tag{3.5}$$

Now move one period back to T_{N-1}, when the second last payment is made. From the vantage point at T_{N-1}, the last payment can be thought of as an option written on the pension asset value. Since the fund follows constant-proportion strategies in the Black-Scholes economy, A_{T_N} follows a lognormal distribution conditional on $A'_{T_{N-1}}$ at time T_{N-1}, i. e.

$$A_{T_N} = A'_{T_{N-1}} \exp\left([\alpha\mu + (1-\alpha)r - \frac{1}{2}\alpha^2\sigma^2](T_N - T_{N-1}) + \alpha\sigma\sqrt{T_N - T_{N-1}}z\right),$$

where z follows the standard normal distribution. Please note that $A'_{T_{N-1}}$ is the pension asset value at T_{N-1} *after* the pension payment, rather than that *before* the payment, $A_{T_{N-1}}$, and that

$$A'_{T_{N-1}} = A_{T_{N-1}} - L^c_{T_{N-1}}.$$

The option value, i. e. $L^f_{T_{N-1}}$, can be obtained using option pricing techniques, for example by solving the Black-Scholes partial differential equation (PDE) of $\pi(\tau, X)$

$$r\pi(\tau, X) = \frac{1}{2}\alpha^2\sigma^2 X^2 \frac{\partial^2\pi}{\partial X^2}(\tau, X) + rX\frac{\partial\pi}{\partial X}(\tau, X) + \frac{\partial\pi}{\partial\tau}(\tau, X) \tag{3.6a}$$

with the boundary condition

$$\pi(T_N, A_{T_N}) = L^c_{T_N}. \tag{3.6b}$$

And $L^f_{T_{N-1}}$ is given by $\pi(T_{N-1}, A'_{T_{N-1}})$. The PDE (3.6) does not have an analytical solution in general. Nevertheless, either through numerical methods like finite different methods, or using an analytical approximation method, one can obtain $L^f_{T_{N-1}}$ as a function of $A'_{T_{N-1}}$, which can also be formulated in terms of $A_{T_{N-1}}$:

$$L^f_{T_{N-1}} = \mathcal{L}^f_{T_{N-1}}(A'_{T_{N-1}}) = \mathcal{L}^f_{T_{N-1}}\left(A_{T_{N-1}} - L^c_{T_{N-1}}\right). \tag{3.7}$$

With the above functional expression for $L^f_{T_{N-1}}$, one can solve for the time-T_{N-1} payment $L^c_{T_{N-1}}$ by applying (3.3) once again:

$$\begin{aligned} L^c_{T_{N-1}} = \min\Bigg\{ \max\Bigg[\frac{1}{2}\Bigg(& L^\ell_{T_{N-1}} - \mathcal{L}^f_{T_{N-1}}\left(A_{T_{N-1}} - L^c_{T_{N-1}}\right) - \beta_{T_{N-1}}K_\ell + \\ & \sqrt{4\beta_{T_{N-1}}A_{T_{N-1}} + \left(\mathcal{L}^f_{T_{N-1}}(A_{T_{N-1}} - L^c_{T_{N-1}}) + L^\ell_{T_{N-1}} - \beta_{T_{N-1}}K_\ell\right)^2}\Bigg), \\ & L^\ell_{T_{N-1}}\Bigg], L^u_{T_{N-1}}\Bigg\} \end{aligned}$$

where $\beta_{T_{N-1}} = \frac{L^u_{T_{N-1}} - L^\ell_{T_{N-1}}}{K_u - K_\ell}$. The above equation characterizes the circularity problem at time T_1 stemming from the conditional indexation. Assuming unique solvability of the above equation,[2] one can, through numerical methods, solve the equation to obtain $L^c_{T_{N-1}}$ as a function of $A_{T_{N-1}}$:

$$L^c_{T_{N-1}} := \mathcal{L}^c_{T_{N-1}}(A_{T_{N-1}}). \tag{3.8}$$

The total liability value at time T_{N-1} is given by

$$L_{T_{N-1}} = L^c_{T_{N-1}} + L^f_{T_{N-1}} = \mathcal{L}^c_{T_{N-1}}(A_{T_{N-1}}) + \mathcal{L}^f_{T_{N-1}}\left(A_{T_{N-1}} - L^c_{T_{N-1}}\right).$$

So one can also write $L_{T_{N-1}}$ as a function of $A_{T_{N-1}}$:

$$L_{T_{N-1}} := \mathcal{L}_{T_{N-1}}(A_{T_{N-1}}) \tag{3.9}$$

Thus we succeed in moving one period back from time T_N to T_{N-1}. Repeating this procedure, we can move back period by period till time T_1, and the total liability value at time T_1 is written as

$$L_{T_1} := \mathcal{L}_{T_1}(A_{T_1}) \tag{3.10}$$

[2]Unique solvability of the circularity equations at every step of the recursion is not easy to verify, since the involved functions are defined implicitly. We take the unique solvability as an assumption here.

Therefore we can come back to the current time t. For the current liability value L_t, one can, through numerical methods, solve the PDE (3.6a) with the boundary condition $\pi(T_1, A_{T_1}) = \mathcal{L}_{T_1}(A_{T_1})$. Finally we arrive at the current liability value and funding ratio:

$$L_t = \pi(t, A_t), \quad FR_t = \frac{A_t}{\pi(t, A_t)}. \tag{3.11}$$

We summarize the computational procedure. The recursion starts from the terminal time T_N, and obtains the amount of the last payment as a function of the asset value as Equation (3.5) by solving the circularity problem at that time point. Then move one period back to the second last payment date T_{N-1}. Using options pricing techniques, the value at T_{N-1} of the last payment can be formulated in the form of (3.7), with which the amount of the first payment can be written as a function of the asset value as Equation (3.8) by solving the circularity problem at T_{N-1}. Thus the total liability value as a function of the asset value is given by Equation (3.9). Following the same procedure that moves from time T_N to T_{N-1}, we can move period by period back till time T_1, and express the time T_1 liability value as (3.10). Moving back the valuation date t, and applying options pricing techniques, we eventually obtain the liability value and the funding ratio as in Equation (3.11).

The cumulative case

The practice that indexation is accumulated over time makes the *minimum* pension benefit paid by pension funds in certain year dependent on the indexation decisions made in previous years, or equivalently on the pension benefit actually paid in previous years. In the multiple-period model, we formulate this type of intertemporal dependence through letting the lower bound of the pension payment in a certain period depend on the *actual* payment in the previous period. In particular, the lower and upper bounds of the pension payment are defined recursively as follows,

$$\begin{aligned}
&L^{\ell}_{T_1} = G, \quad L^{u}_{T_1} = G\exp[\rho(T_1 - T_0)] \\
&L^{\ell}_{T_i} = L^{c}_{T_{i-1}}, \quad L^{u}_{T_i} = L^{c}_{T_{i-1}}\exp[\rho(T_i - T_{i-1})], \quad \text{for} \quad i = 2, \ldots, N.
\end{aligned}$$

Because of the recursive nature of the pension payment, the pension payment in a certain period is dependent on the actual payments in previous periods. Thus additional state variables need to be introduced to address the added intertemporal dependence arising from the cumulative case. The required number of additional state variables depends on the structure of the pension fund, especially the number of generations of participants in the fund. For ease of exposition, we distinguish a single-generation fund and a multiple-generation fund, and address the single-generation fund first.

Consider the pension fund with a single generation. The computational procedure also works backwards and starts from the last payment date T_N. The solution of the circularity problem at time T_N is given by

$$L^c_{T_N} = \min\left\{\max\left[\frac{1}{2}\left(L^\ell_{T_N} - \beta_{T_N}K_\ell + \sqrt{4\beta_{T_N}A_{T_N} + (L^\ell_{T_N} - \beta_{T_N}K_\ell)^2}\right), L^\ell_{T_N}\right], L^u_{T_N}\right\},$$

where $\beta_{T_N} = \frac{L^u_{T_N} - L^\ell_{T_N}}{K_u - K_\ell}$. Because $L^\ell_{T_N}$ and $L^u_{T_N}$ are dependent on the actual payment in the previous period, $L^c_{T_{N-1}}$, the pension payment at time T_N can then be written as

$$L^c_{T_N} = \mathcal{L}^c_{T_N}(A_{T_N}, L^c_{T_{N-1}}). \tag{3.12}$$

Observe that in comparison to the corresponding equation (3.5) in the noncumulative case, Equation (3.12) includes the actual payment in the previous period as an additional state variable.

Now move one period back to T_{N-1}, when the second last payment is made. In principle, one can apply the options pricing techniques to obtain $L^f_{T_{N-1}}$ as a function of $L^c_{T_{N-1}}$ and $A'_{T_{N-1}}$, which can also be formulated in terms of $A_{T_{N-1}}$. That is,

$$\begin{aligned} L^f_{T_{N-1}} &= \mathcal{L}^f_{T_{N-1}}(A'_{T_{N-1}}, L^c_{T_{N-1}}) \\ &= \mathcal{L}^f_{T_{N-1}}\left(A_{T_{N-1}} - L^c_{T_{N-1}}, L^c_{T_{N-1}}\right). \end{aligned} \tag{3.13}$$

With the above functional expression for $L^f_{T_{N-1}}$, one can solve for the second last payment $L^c_{T_{N-1}}$ through solving the circularity problem at T_{N-1} by means of (3.3):

$$L^c_{T_{N-1}} = \min\left\{\max\left[\frac{1}{2}\left(L^\ell_{T_{N-1}} - \mathcal{L}^f_{T_{N-1}}\left(A_{T_{N-1}} - L^c_{T_{N-1}}, L^c_{T_{N-1}}\right) - \beta_{T_{N-1}}K_\ell + \sqrt{4\beta_{T_{N-1}}A_{T_{N-1}} + \left(\mathcal{L}^f_{T_{N-1}}\left(A_{T_{N-1}} - L^c_{T_{N-1}}, L^c_{T_{N-1}}\right) + L^\ell_{T_{N-1}} - \beta_{T_{N-1}}K_\ell\right)^2}\right), L^\ell_{T_{N-1}}\right], L^u_{T_{N-1}}\right\}$$

where $\beta_{T_{N-1}} = \frac{L^u_{T_{N-1}} - L^\ell_{T_{N-1}}}{K_u - K_\ell}$. Assuming unique solvability of the above equation, one can solve the equation to obtain $L^c_{T_{N-1}}$ as a function of the following form:

$$L^c_{T_{N-1}} := \mathcal{L}^c_{T_{N-1}}(A_{T_{N-1}}, L^c_{T_{N-2}}), \tag{3.14}$$

for the reason that the lower and upper boundaries of pension payment at time T_{N-1}, $L^\ell_{T_{N-1}}$ and $L^u_{T_{N-1}}$, are dependent on the actual payment at time T_{N-2}, $L^c_{T_{N-2}}$. As such, the total liability value at time T_1 is computed as

$$\begin{aligned} L_{T_{N-1}} &= L^c_{T_{N-1}} + L^f_{T_{N-1}} \\ &= \mathcal{L}^c_{T_{N-1}}(A_{T_{N-1}}, L^c_{T_{N-2}}) + \mathcal{L}^f_{T_{N-1}}\left(A_{T_{N-1}} - L^c_{T_{N-1}}, L^c_{T_{N-1}}\right). \end{aligned}$$

So one can also write L_{T_1} as

$$L_{T_{N-1}} := \mathcal{L}^c_{T_{N-1}}(A_{T_{N-1}}, L^c_{T_{N-2}}) \tag{3.15}$$

Repeating this procedure, we can move back period by period till time T_2, when the second payment is made, and the total liability value at time T_2 is written as

$$L_{T_2} := \mathcal{L}_{T_2}(A_{T_2}, L^c_{T_1}).$$

Then we move one period back from T_2 to T_1. Note that the lower and upper boundaries of pension payment at time T_1, $L^\ell_{T_1}$ and $L^u_{T_1}$, are known parameters rather than unknown variables. Therefore the recursion from T_2 to T_1 leads to the time-T_1 liability value given by

$$L_{T_1} := \mathcal{L}_{T_1}(A_{T_1}). \tag{3.16}$$

At last we can come back to the current time t. For the current liability value L_t, one can, through numerical methods, solve the PDE (3.6a) with the boundary condition $\pi(T_1, A_{T_1}) = \mathcal{L}_{T_1}(A_{T_1})$. Finally we arrive at the current liability value and funding ratio:

$$L_t = \pi(t, A_t), \quad FR_t = \frac{A_t}{\pi(t, A_t)}. \tag{3.17}$$

In comparison with the noncumulative case, the one-generation pension fund with cumulative indexation requires one additional state variable for the computation. This computational procedure can be generalized to a multiple-generation fund with cumulative indexation by including the current pension payments (or the current promised payments) of every generation of the participants as state variables. In the cumulative case of a multiple-generation fund, however, the required number of state variables increases with the number of generations of participants in the fund, in order to carry relevant information about the different indexation applied to different generations. A realistic model of the cumulative case needs a large number of state variables (up to 80) to carry the relevant information on indexation decisions over time.

With such realistic modeling of the cumulative case, the valuation of conditionally indexed pension liabilities in the consistent implementation leads to a computational task of high dimension. The model can be solved in principle through backward recursion as in our stylized model, but a major computational problem is the dimension of the state space which may be beyond the computational capacity usually available. The dimensionality problem cannot be solved by standard Monte Carlo methods since the liability value is needed to determine what indexation decisions are made; adaptations akin to Monte Carlo methods for American option pricing may be possible but are not addressed in this chapter.

3.3 Numerical illustration

In this section, we present some numerical exercises in the cumulative case. We consider the simplest case where there are two pension payments ($N = 2$). In the numerical illustration, the following parameter values are assumed:

$$
\begin{aligned}
& r = 3\%, \quad \mu = 7\%, \quad \sigma = 20\%, \quad \rho = 4\%, \quad \alpha \in \{0.25, 0.50, 0.75\} \\
& T_0 = 0, \quad t = 9, \quad T_1 = 10, \quad T_2 = 20 \\
& K_\ell = 110\%, \qquad K_u \in \{115\%, 140\%, 160\%\} \\
& L^\ell_{T_1} = 100, \qquad L^u_{T_1} = L^\ell_{T_1} e^{\rho(T_1 - T_0)} = 149
\end{aligned}
$$

3.3.1 The impact of investment strategies

To see the impact of the aggressiveness of investment strategies, we fix the policy ladder ($K_\ell = 110\%$ and $K_u = 140\%$), and consider three stock weights $\alpha \in \{25\%, 50\%, 75\%\}$.

Before looking into the liability value and funding ratio of the conditional indexation schemes, let us recall some defining properties of defined-benefits (DB) and defined-contribution (DC) schemes as reflected by liability value and funding ratio. For a pure DB scheme, because of the bond-like property of pension benefits, the liability value is

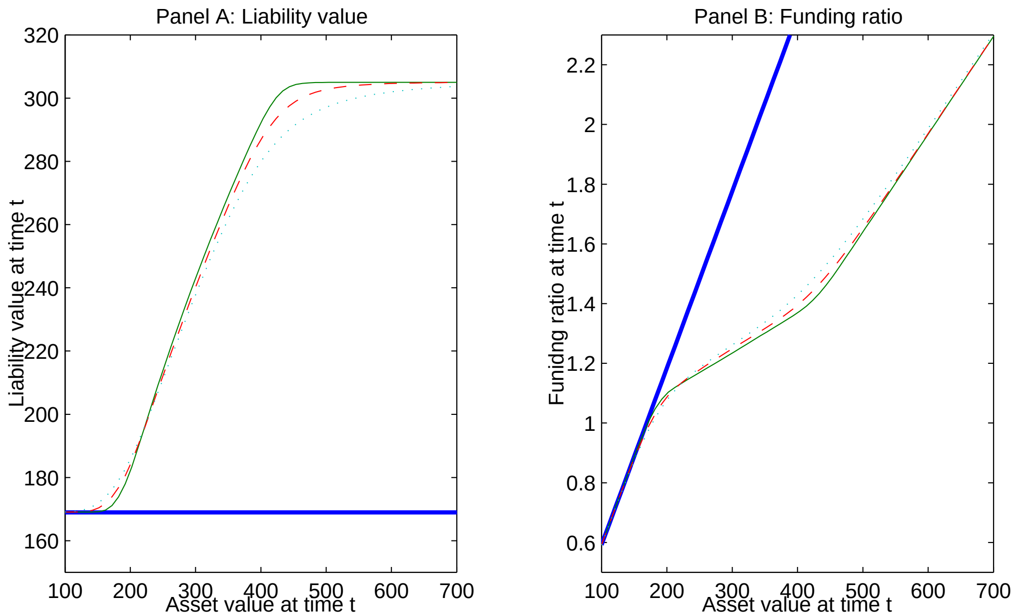

Figure 3.1: **Liability value and funding ratio at time t for different investment policies** This figure shows both the liability value (the left panel) and the funding ratio (the right panel) at time t as a function of the then asset value for three asset mixes: the stock weight is assumed to be 25% (the solid line), 50% (the dashed line), and 75% (the dotted line). The bold lines are for the financial measure on the basis of the FTK requirement, which will be discussed in Section 3.4.

independent of changes of pension asset value, and the funding ratio is therefore a linear function of the pension asset value. On the other hand, for a typical DC scheme, the liability value is equal to (i. e. a linear function of) the pension asset value whereas the funding ratio is constant at 100%.

As can be seen from the behavior of its liability value and funding ratio illustrated in Figure 3.1, conditional indexation schemes are neither typical of DB nor of DC, but strike some balance between the two stereotypes. If the asset value is either very low or very high, the liability values for all the three different investment policy show little or even no dependence of the asset value, bearing the hallmark of DB schemes. The same point can be seen from the funding ratio of the conditional indexation scheme as it approaches a linear function of the asset value for both ends of the asset value domain as shown in the figure. In contrast, for the intermediate domain of asset values, say between 200 and 400, the conditional indexation scheme bears a close resemblance to DC schemes, in that the liability value is increasing with the asset value, and the funding ratio is insensitive to the variation of asset value.

Therefore the practice of conditional indexation has the effect of stabilizing the funding ratio through introducing a DC element to the originally DB system. The magnitude

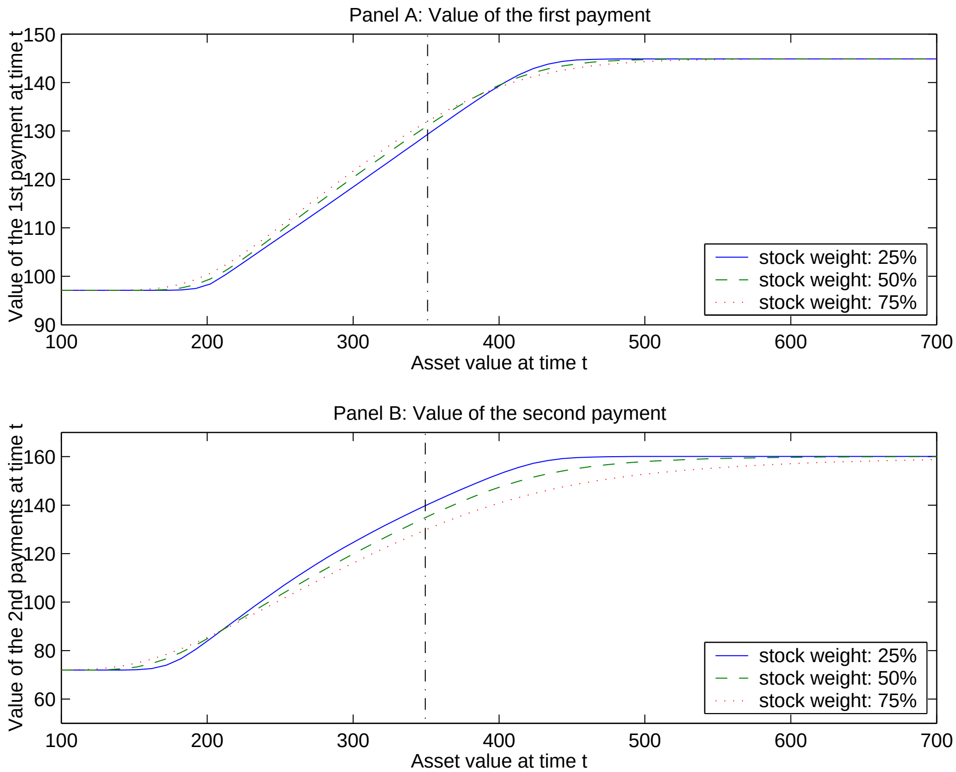

Figure 3.2: **Value of the two pension payments to be made at T_1 and T_2** This figure shows the liability value of the first payment(the upper panel) and the second payment (the lower panel) at time t as a function of the asset value for three asset mixes: the stock weight is assumed to be 25% (the solid line), 50% (the dashed line), and 75% (the dotted line). The vertical dash-dotted line is corresponding to asset value equal to 350.

of the stabilizing effect, however, depends on the investment policy. The stability of funding ratio in the intermediate region of asset values becomes more pronounced with a lower stock weight, for in that case the curve for the funding ratio is flatter as shown in the figure. An intuitive explanation is as follows. In comparison with a scheme with a high stock weight, a low stock weight scheme has less volatility of asset value. For a conservative asset mix, a given amount of asset value increase implies an increased future asset value with more certainty than for an aggressive asset mix. And hence a larger pension increase in the form of an increased indexation level can be granted to participants, which implies a larger increase in pension liabilities. That is, the liability value of is more sensitive to the variation of asset value for a conservative asset mix. Translated into the funding ratio, it implies that the funding ratio has greater stability when the investment strategy is less risky.

Whether increased riskiness of asset mix will raise or lower the liability value depends on the current asset value. In particular, when the asset value is low, say equal to

200, a higher stock weight increases the liability value. In contrast, when the asset value is high, say equal to 400, a higher stock weight decreases the liability value. The dependence on the current asset value stems from the collar structure of pension rights under conditional indexation. A higher stock weight may enhance the probability of granting greater-than-the-minimum indexation, and *increase* the liability value. On the other hand, it is also possible that an increased stock weight dampens the possibility of the maximum indexation, and hence *decreases* the liability value. The balance of the two countervailing effects depends on the current asset value. The increasing effect dominates for low asset values due to the guaranteed pension floor, whereas the decreasing effect prevails for high asset values because of the full indexation ceiling of pension rights. At some intermediate values of asset these two effects compensate each other, making the liability value insusceptible to changes of asset mix. The impact of investment policy can also been expressed in terms of the funding ratio: a raised stock weight decreases the funding ratio at low asset values, and increase the funding ratio at high asset values.

The pension liability consists of the two payments to be made in the future in this example, and the liability value can be decomposed into two components: the value of the early payment, and that of the late payment. To analyze the impact of investment policy on the two components can lead one to see how the change of investment policy results in *intertemporal* redistribution. Figure 3.2 shows the values of the two payments as a function of the asset value for the three investment policies. When the current asset is either very high or very low, say 200 and 450, the change of asset mix drives the two components in the same direction. In contrast, for some intermediate asset values, the change of asset mix decreases the value of one payment, but increases the value of another. For instance, consider the asset value equal to 350, where a rise in stock weight will increase the value of the first payment, but decrease the value of the second considerably. That is to say, the improved growth potential from higher stock weight is captured by the first payment preemptively, leaving the second payment only a loss. If the first payment is interpreted as the pension rights of an old generation, and the second as that of a young generation, then it implies that the investment policy has implications for *inter-generational* redistribution.

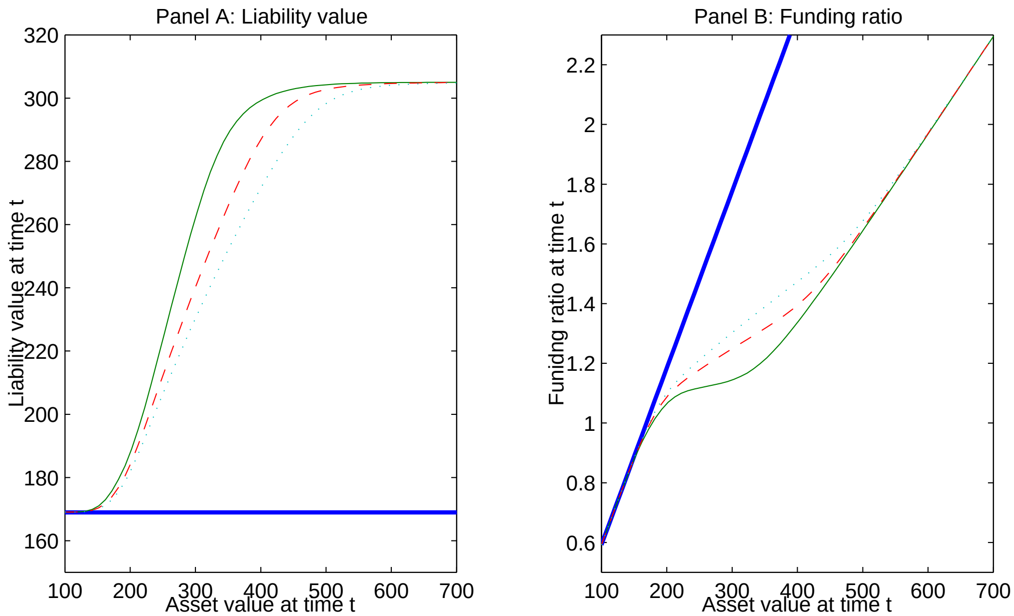

Figure 3.3: **Liability value and funding ratio at time t for different policy ladders**
This figure shows both the liability value (the left panel) and the funding ratio (the right panel) at time t as a function of the then asset value for three policy ladders: K_u is assumed to be 1.15 (the solid line), 1.40 (the dashed line), and 1.60 (the dotted line). The bold lines are on the basis of valuation following the FTK requirement, which will be discussed in Section 3.4.

3.3.2 The impact of policy ladders

To see the impact of the policy ladders, we fix the stock weight ($\alpha = 0.5$) and vary the threshold of funding ratio for full indexation: $K_u \in \{115\%, 140\%, 160\%\}$. To the extent that a lower value of K_u implies that it is easier for participants to receive full indexation, the three values for K_u are used to characterize how generous the policy ladder is.

As would be anticipated, a more generous policy ladder, *ceteris paribus*, results in a greater liability value and hence a lower funding ratio whatever the current asset value is (Figure 3.3). Another interesting point is that the funding ratio is more stable under a more generous policy ladder as the curve of funding ratio is flatter in such a case. An intuitive explanation is that a more generous policy ladder makes the granted indexation level and hence the liability value more responsive to the variation of asset value, which implies more stability of the funding ratio.

As Figure 3.4 shows, the values of both payments are increasing with the generosity of policy ladder, irrespective of the current asset value. However, generally the first payment seems to benefit more from a generous policy ladder than the second does, reflecting the preemptive advantage of the first payment over the second.

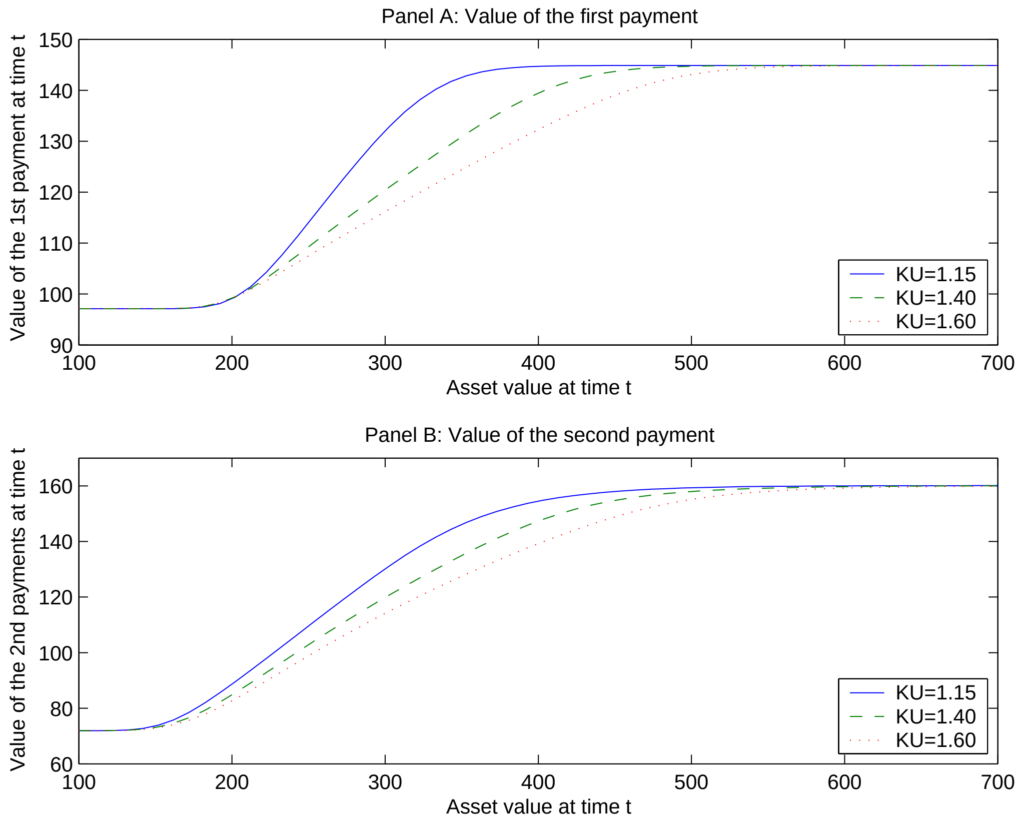

Figure 3.4: **Liability value and funding ratio at time t for different policy ladders** This figure shows the liability value of the first payment(the upper panel) and the second payment (the lower panel) at time t as a function of the asset value for three asset mixes: K_u is assumed to be 1.15 (the solid line), 1.40 (the dashed line), and 1.60 (the dotted line).

3.4 Comparison to the proxy-based implementation

In the consistent implementation discussed above, the funding ratio used for indexation decisions is financially valid, and hence it is a reliable indicator in assessing the financial status of the pension fund. (The funding ratio obtained in this way will be referred to as the "consistent funding ratio" in the following.) In the valuation of pension liabilities, the circularity problem needs to solved to keep the consistency of the resultant funding ratio between its role as the basis of indexation and its role as an indicator of financial solvency. The current practice of conditional indexation is different though. Indexation decisions are made on the basis of a proxy of funding ratio for which the liability value is computed assuming a fixed indexation, typically zero indexation.[3] The proxy-based implementation avoids the circularity problem inherent in the consistent implementation.

[3]According to the current Dutch pension regulation (FTK), the liability value is computed on the basis of the guaranteed pension rights only, excluding the contingent rights from conditional indexation. See http://docs.szw.nl/pdf/35/2004/35_2004_349957.pdf (in Dutch).

A caveat for the current practice is that the zero-indexation proxy used for indexation decisions may be misleading in assessing the financial status of pension funds. From the perspective of market valuation, it is straightforward that the zero-indexation proxy overstates the financial solvency of pension funds, and hence is an overestimate of the "actual" funding ratio, as long as it is possible to grant indexation to participants. Nevertheless, one may defend the validity of the zero-indexation proxy as an indicator of financial solvency by the argument that the proxy, if being read with some "discount", can still make sense. As will be shown below, however, the proxy overestimates the latent actual funding ratio in a complicated manner, making such discounting rather difficult. Moreover, the funding ratio computed on the basis of *any* fixed level of indexation is not a reliable indicator of financial soundness, because the indexation levels decided in the future are contingent and varied. Therefore for the purpose of assessing the financial status of pension funds using this implementation, the actual funding ratio needs to be computed and used. In the following, we shall address the market valuation of the liability and the computation of the actual funding ratio in the proxy-based implementation.

A different funding ratio used for indexation decisions, *ceteris paribus*, leads to different indexation decisions and hence to a different pension profile. The zero-indexation proxy is of course different from the consistent funding ratio, so the scheme resulting from zero-indexation proxy, referred to as the the "ZIP scheme" hereafter, is different from that generated by the consistent implementation. The question we face is what the actual funding ratio of the ZIP scheme is. It is worth noting that the consistent funding ratio is *not* the answer since it is the actual funding ratio for a *different* pension scheme.

To obtain the actual funding ratio of the ZIP scheme, we need to know the zero-indexation proxy first, for the reason that it is on the basis of the zero-indexation proxy that indexation decisions and pension rights are determined. We still work in the two-payment model. With the assumption of zero-indexation, the (proxy) liability value at t is simply the two minimum payment at T_1 and T_2 discounted by the risk-free rate: $L_t^{zip} = L_\ell^{T_1} e^{-r(T_1-t)} + L_\ell^{T_1} e^{-r(T_2-t)}$. Therefore the zero-indexation proxy is

$$FR_t^{zip} = \frac{A_t}{L_t^{zip}}. \tag{3.18}$$

Their values as a function of the asset value are shown in Figure 3.1 and 3.3. They are of DB nature, and independent of investment policy and policy ladder. We assume that the two indexation decisions at T_1 and T_2 are based on the two zero-indexation proxies immediately before payments respectively. The zero-indexation proxies at T_1 and T_2 can be computed in a similar way.

Given the way the indexation decisions are made and the way the proxies for the indexation decisions are computed, one can obtain the actual funding ratio of ZIP schemes by applying classical options pricing techniques. Key to the computation of the funding ratio is the market valuation of pension liabilities. In this context, the valuation of liabilities is to price two contingent pension payments at T_1 and T_2. Applying Monte Carlo simulation methods, one can obtain the market value of the liability and the actual funding ratio of the ZIP scheme.

Using the same parameter values as before, the upper panels in Figures 3.5 and 3.6 show the relation between the actual funding and the zero-indexation proxy of the ZIP scheme. Because the zero-indexation proxy, as can be seen from (3.18), is a linear function of the asset value, these two figures can also be interpreted as representing the actual funding ratio as a function of the asset value. A striking point, as can be seen from the downward-sloping section of some curves in the figures, is that an increase in the zero-indexation proxy reduces the actual funding ratio for some investment strategies and policy ladders. It is because an increase in the asset value, *ceteris paribus*, would raise the first payment, which in turn raise the lower bound of the second payment, and the boosting effect on both payments can raise the liability value to such an extent that is greater than the increase in the asset value. It makes no sense, based solely on the numerical exercises of a stylized model as used in this chapter, to rush to the conclusion that a stock market crash and a dampened asset value should be welcomed with open arms by pensions funds that wish to improve their financial status. Nevertheless, it highlights the fact that it is possible for an increase in the asset value to in effect hurt the actual funding ratio of a pension fund.

The existence of such a counterintuitive scenario is accounted for by two elements inherent in the proxy-based implementation. One is the dynamic inconsistency of this implementation. The very fundamental idea of conditional indexation is that a pension fund pays more only when it possesses more vis-à-vis its liability. The proxy-based implementation is not able to ensure that this idea is realized, because a proxy of funding ratio simply uses a “financially incorrect” measure of the liability, and hence cannot tell one whether a pension fund possesses more relative to its liability. There may exist the case where a pension fund pays more when it actually possesses less relative to its

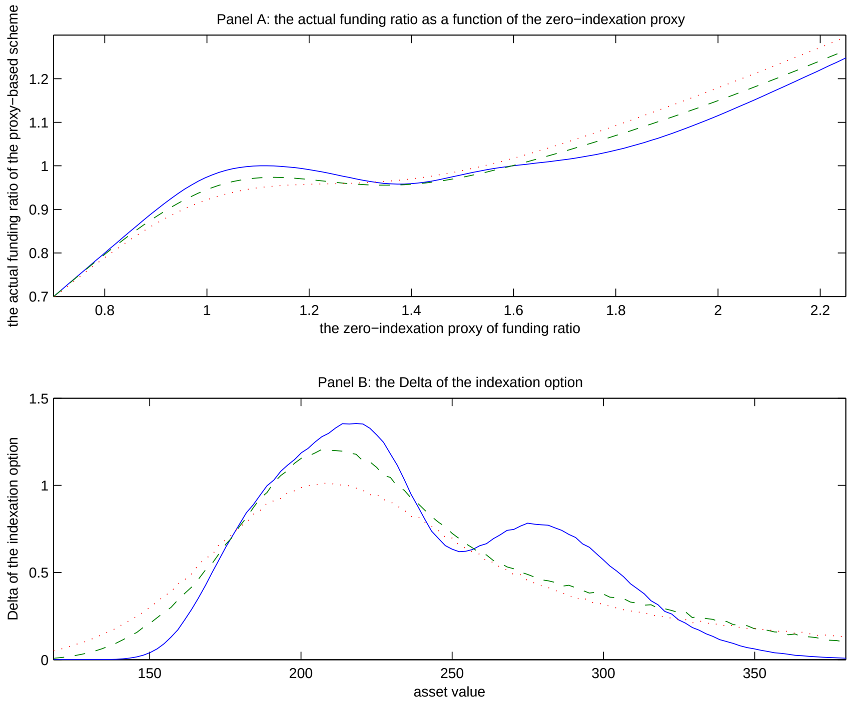

Figure 3.5: **The proxy-based implementation for different stock weights** In this figure, Panel A shows the actual funding ratio at time t as a function of the zero-indexation proxy, and Panel B shows the first-order derivative of the actual liability value with respect to the asset value (the Delta of the indexation option). The stock weight is assumed to be 25% (the solid line), 50% (the dashed line), and 75% (the dotted line). The horizontal axes of both panels are set to correspond to each other.

zero-indexation proxy		**1.00**	**1.10**	**1.20**	**1.40**	**1.60**	**1.80**
Panel A investment policy: $\alpha = \cdots$	25%	*0.97*	*1.00*	*0.99*	*0.96*	*1.00*	*1.04*
		`0.99`	`1.06`	`1.10`	`1.15`	`1.20`	`1.24`
	50%	*0.95*	*0.97*	*0.97*	*0.96*	*1.00*	*1.07*
		`0.97`	`1.04`	`1.09`	`1.16`	`1.21`	`1.25`
	75%	*0.92*	*0.95*	*0.96*	*0.97*	*1.02*	*1.09*
		`0.96`	`1.03`	`1.08`	`1.16`	`1.22`	`1.27`
Panel B policy ladder: $K_u = \cdots$	115%	*0.91*	*0.89*	*0.86*	*0.91*	*0.98*	*1.05*
		`0.97`	`1.03`	`1.07`	`1.11`	`1.13`	`1.15`
	140%	*0.95*	*0.97*	*0.97*	*0.96*	*1.00*	*1.07*
		`0.97`	`1.04`	`1.09`	`1.16`	`1.21`	`1.25`
	160%	*0.96*	*1.00*	*1.02*	*1.02*	*1.04*	*1.09*
		`0.98`	`1.05`	`1.11`	`1.19`	`1.25`	`1.31`

Table 3.1: **Comparison of funding ratios** The table shows the actual funding ratio of the ZIP scheme and the consistent funding ratio corresponding to various values of the zero-indexation proxy. Panel A is for three stock weights with a fixed policy ladder ($K_\ell = 110\%$, and $K_u = 140\%$) whereas Panel B is for three policy ladders with a fixed stock weight ($\alpha = 50\%$). The actual funding ratios of the ZIP scheme are rendered in *italics*, and the consistent funding ratios in `typewriter` font.

liability. The other element contributing to this scenario is the accumulative practice of indexation. It allows an indexation decision in a year to have "repercussions" on the pension benefit over many years that follows. For instance, if a participant receives 2% indexation at the year of retirement, her pension benefit in every year till her decease will be raised by 2% compared to the case of no indexation. Paying 2% more in one year is no big deal, but an increase of 2% in all 30 annual payments may have noticeable impact on the liability value.

Therefore it is possible for a pension fund to grant participants too much indexation in the sense that an increase in asset value leads to a lower actual funding ratio. The possibility of over-indexation can also be seen from the Delta of the indexation option, i. e. the first-order derivative of the liability value as a function of the asset value. When the Delta is greater than one, say 1.5, it means that an increase of 1 dollar in the asset value leads to an increase of 1.5 dollars in the liability value stemming from more indexation. As shown in the lower panels in Figures 3.5 and 3.6, the Delta is indeed larger than one for some combinations of the current asset value, the stock weight, and the policy ladder, corresponding to what we see from the behavior of the actual funding ratio in the corresponding upper panels. It is intuitively clear that a more generous indexation rule is more likely to have over-indexation (Figure 3.6). In a less intuitive way, Figure 3.5 implies that a pension fund with a more aggressive investment strategy is less likely to have over-indexation.

Also evident in the upper panels in Figures 3.5 and 3.6 is that the actual funding ratio of the ZIP scheme is insensitive to the change of the asset value, a phenomenon also seen in the consistent implementation. Akin to the case of the consistent implementation, the actual funding ratio as a function of the asset value depends on the investment strategy and the policy ladder chosen by the pension fund. Compare Figures 3.5 with Panel B of Figure 3.1, and compare Figure 3.6 with Panel B of Figure 3.3. One can find that the pattern of the dependence on the investment strategy and on the policy ladder is similar to the case in the consistent implementation. The intuitive explanation we offered in the case of the consistent implementation seems also plausible here.

Given the relevance of these various factors, it is possible that the actual funding ratio of Fund A is greater than that of Fund B, but the zero-indexation proxies of the two funds indicate the other way around. It implies that the comparison of financial solvency of different funds based solely on their zero-indexation proxies can be misleading. Furthermore, even for a single fund, an improved zero-indexation proxy may not indicate improved financial solvency if there are major changes in relevant aspects. The complicated relation between the zero-indexation proxy and the actual funding ratios implies there does not exist a simple way to "discount" the zero-indexation proxy to learn about the actual financial status of pension funds. Ideally, reliable discounting should take into account current asset value, investment policy, policy ladder, and demographic composition of pension funds.

Table 3.1 compares the three funding ratios discussed in this chapter: the consistent funding ratio, the zero-indexation proxy, and the actual funding ratio of the ZIP scheme. For a given asset value, the zero-indexation proxy is greater than the consistent funding ratio, which in turn is greater than the actual funding ratio of the ZIP scheme. So it is predictable that a change from the implementation based on zero-indexation proxy to the consistent implementation, *ceteris paribus*, will improve the financial soundness of pension funds. The reason for the improvement is that after the change of valuation method, indexation decisions are made based on less favorable (to participants) funding ratios, resulting in less indexation, and thus less pension liabilities.

3.5 Conclusion

In this chapter, we addressed the market valuation of conditionally indexed pension liabilities in two types of implementation. One is the current practice, where a proxy of funding ratio, typically the zero-indexation proxy, is used for indexation decisions. The other is to use the market-based and financially correct funding ratio for indexation decisions. Given its conformity with the ongoing move to market-based accounting standards, the latter implementation, known as the consistent implementation here, may be able to find its way into practice; the valuation of conditionally granted pension liability is on the agenda of the pension regulator in the Netherlands,[4] where many pension funds have adopted conditional indexation.

In this chapter, we developed a backward recursion approach to the liability valuation in the consistent implementation. Based on the same model, we also considered the liability valuation in the proxy-based implementation. Numerical examples show that in both types of implementation, pension funds with conditional indexation find a middle

[4] See "Principles for a financial assessment framework: more transparency and clearer information" by the Dutch central bank, DNB.

road between DB and DC to the extent that the financial status depends on the asset value of pension funds. Conditional indexation provides a shield of the financial solvency against the fluctuation of the asset value through introducing a DC element into the originally DB system. The numerical examples also highlight the impact on the financial status of the investment strategy and the indexation rule that a pension fund adopts.

In the proxy-based implementation, the actual funding ratio, as opposed to the proxy, needs to be computed and used in assessing the financial soundness of pension funds. Some numerical exercises show that in this respect, the zero-indexation proxy may be misleading: an increase in the zero-indexation proxy (or equivalently in the asset value) may hurt the actual funding ratio of a pension fund. The actual funding ratio is directly related to the distribution of future actual funding ratios, which form the basis of pension solvency regulation. The existence of such peculiar scenarios is related to the dynamic inconsistency inherent in this type of implementation. Moreover, the comparison of the financial status among pension funds based solely on the proxy is rendered invalid by the fact that the relation between the proxy and the actual funding ratio depends on the investment policy, the indexation rule and the demographics of pension funds.

In the consistent implementation, the cumulative practice of indexation leads to a computational problem of high dimension if a more realistic model with multiple generations is adopted. The dimensionality problem is not addressed here, and invites further research. Other interesting avenues for future research include examining the extent to which the regulation using the zero-indexation proxy is suboptimal, and conducting an welfare analysis of the pension contracts resulting from implementing conditional indexation in different ways.

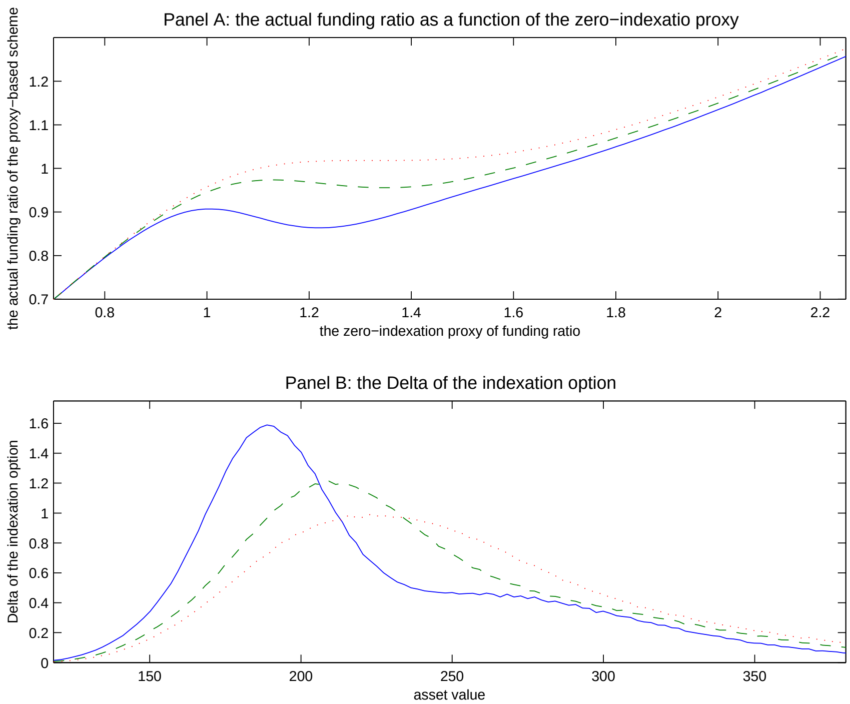

Figure 3.6: **The proxy-based implementation for different policy ladders** In this figure, Panel A shows the actual funding ratio at time t as a function of the zero-indexation proxy, and Panel B shows the first-order derivative of the actual liability value with respect to the asset value (the Delta of the indexation option). The upper threshold of funding ratio K_u is assumed to be 1.15 (the solid line), 1.40 (the dashed line), and 1.60 (the dotted line). The horizontal axes of both panels are set to correspond to each other.

Chapter 4

Portfolio Choices and Consumption Smoothing under Time-variant Equity Premia and State Uncertainty

4.1 Introduction

There has been well documented empirical evidence that the equity premia are time-varying rather than constant (see, for instance, Poterba and Summers [1988], Fama and French [1989], and Cochrane [2005]). Expected returns on common stocks have been found to vary over business cycles, and the general message of the empirical evidence is that expected returns are lower when business conditions are strong and higher when business conditions are weak. The implications of time-variation in expected returns are widely recognized by both academics and investment professionals. Some recent papers, for instance Kim and Omberg [1996] and Wachter [2002], have studied the implications for dynamic asset allocation of this deviation from the traditional view of constant expected returns. In this situation, optimal consumption and investment strategies involve timing business conditions. Optimal strategies typically require more allocation to stocks when business conditions are relatively weak.

It has long been recognized that investors do not have complete information about business conditions or the state of economy. Institutional and individual investors usually disagree on the assessment of the current business conditions, and economic prospects. This disagreement is symptomatic of the fact that there is quite some uncertainty surrounding the state of the economy. In the context of time-varying expected returns, state uncertainty leads to uncertainty about expected returns, which an investor must take into account in choosing the optimal consumption and investment. Extensive work has been undertaken on the implications of state uncertainty for portfolio analysis. In a continuous-time setting, Detemple [1986], Dothan and Feldman [1986], and Gennotte [1986] examine the portfolio problem of an investor who cannot observe the true state of the economy. An important finding of these authors is that the investor's decision problem in the presence of unobservable state variables may be decomposed into two separate problems: an inference problem in which the investor estimates the current values of the state variables, and an investment optimization problem which is solved by treating the estimated values of the state variable as state variables themselves.

In this chapter, we examine the optimal consumption and asset allocation in a setting where expected returns on stocks are time-varying, but unobservable. In particular, the unobservable expected return is assumed to follow a mean-reverting process. The model can be viewed as an extension of Kim and Omberg [1996] by assuming unobservability of the expected return. Using the estimation-optimization procedure leads to the solution of the problem. In the presence of state uncertainty, the investor's assessment of the investment opportunity set is imperfect in the sense that the estimation error does not vanish over time. Moreover, after observing a sufficiently long history of realized returns on stocks, the investor can no longer improve upon the estimation error. We refer to the state where the estimation error remains unchanged over time as the stable phase. We find that in the stable phase, the optimal consumption and asset allocation can be solved in closed form, which provides insights into the consumption and investment behavior in the presence of state uncertainty.

This chapter is related to the work by Brennan [1998], who analyzes the dynamic portfolio problem of an investor who knows the expected return is constant but is uncertain about its value. The focus of Brennan's study is on the role of future learning about the expected return in asset allocation, and his numerical exercise shows state uncertainty has a significant effect on portfolio choice. Our study can be viewed as an extension of Brennan [1998] in that when the mean-reverting characterization of the expected return is reduced to a constant by specific choice of parameter values, our model corresponds to his constant expected return case. Note that if the expected return is constant, the estimation error will vanish as the observation interval increases. Thus state uncertainty does not affect portfolio decisions in the long run. In contrast, when there is time variation in the expected return, the estimation error always exists and affects investor's choice.

In this study, the investor is allowed to derive utility from terminal wealth as well as intermediate consumption. When this model is applied to provide insights to the financial decisions of collective defined-contribution pension funds, the consumption plan in the model can be interpreted as the benefit policy of pension funds. Because the benefit policy of pension funds has been a subject of heated debate in recent years, the implications of consumption smoothing arising from this study for the benefit policy are discussed.

The remainder of the chapter is organized as follows. The model and the general solution method are presented in Section 4.2. Section 4.3 is devoted to the optimal strategy in the stable phase where the investor can no longer improve upon the estimation error. In section 5.5, we offer some representative numerical examples and discussion. Section 5.6 concludes.

4.2 The investor's decision problem

Consider an investor who is concerned with maximizing the expected utility defined over lifetime consumption and wealth at the end of a horizon, T. The investor can trade continuously in a riskless bond and a risky stock. For the riskless bond, the constant interest rate is denoted by r, so that the price of the bond at time t is described by

$$dB_t = rB_t dt. \tag{4.1}$$

The stock price S_t is characterized by the following stochastic differential equation:

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_1 dZ_{S,t}, \tag{4.2}$$

where $Z_{S,t}$ is a standard Brownian motion and σ_1 is a positive constant. μ_t is the expected return of the stock, and assumed to follow an Ornstein-Uhlenbeck process:

$$d\mu_t = \theta(\bar{\mu} - \mu_t)dt + \sigma_2 dZ_{\mu,t}, \tag{4.3}$$

where $Z_{\mu,t}$ is another standard Brownian motion jointly normally distributed with $Z_{S,t}$, and $\bar{\mu}$, θ, and σ_2 are positive constants. This characterization of financial markets reflects the empirical finding that equity markets have time-varying risk premia, and it is adopted by Kim and Omberg [1996] in their investigation of optimal portfolio decisions. What distinguishes this study from that of Kim and Omberg [1996] is that in this study we assume the expected return μ_t to be unobservable to the investor. Whereas all the parameter values are assumed to be known to the investor, the unobservability assumption of the expected return serves to characterize the incomplete information and state uncertainty that are often, if not always, confronting investors.

The assumption that all the parameters are known is a stringent one, since generally investors do not know for sure either the current expected return or the dynamics governing the evolution of the expected return. For a more realistic model, for instance, one may also allow for parameter uncertainty.[1] However, we opt to simplify by modeling the incomplete information as an unobservable variable generated by a known process. This model can be viewed as a generalization of the study by Brennan [1998], because if $\theta = \sigma_2 = 0$, the model corresponds to the problem studied by him.

In the spirit of Fleming and Rishel [1975], Gennotte [1986], Detemple [1986], and Dothan and Feldman [1986], the investor's decision problem in this situation can be decomposed into two separate problems: an inference problem in which the investor estimates the current value of the expected return, μ_t, and an investment problem which is solved by treating the estimated value of the expected return as the expected return itself.

First consider the inference problem. Denote the mean and variance of the estimation of the expected return, μ_t, by m_t and V_t. Assume that at the beginning of the observation interval, the investor views the distribution of μ_0 as a Gaussian distribution with mean m_0 and variance V_0. Then it follows from Gennotte [1986] and Liptser and Shiryayev [1978] that

$$dm_t = \theta(\bar{\mu} - m_t)dt + \frac{\sigma_1\sigma_2\rho + V_t}{\sigma_1^2}\left(\frac{dS_t}{S_t} - m_t dt\right), \tag{4.4}$$

$$dV_t = \left[-2\theta V_t + \sigma_2^2 - \sigma_1^{-2}(\sigma_1\sigma_2\rho + V_t)^2\right] dt, \tag{4.5}$$

where ρ is the correlation coefficient between the two Brownian motions, $Z_{S,t}$ and $Z_{\mu,t}$. Equation (4.4) summarizes the investor's update of the investment opportunity over time. Intuitively, the investor tends to raise her assessment of the mean when it is below the known long-term mean of the expected return ($\bar{\mu} - m_t > 0$), and vice versa. In addition,

[1] Xia [2001] shows that uncertainty of parameters affects the optimal portfolio choice through dynamic learning.

the investor updates her assessment by comparing the security return (dS_t/S_t) and her current assessment of the mean. Assuming $\sigma_1\sigma_2\rho + V_t > 0$, the investor tends to raise her assessment whenever the security return is above her current assessment. When $\sigma_1\sigma_2\rho + V_t < 0$, the opposite is true. Equation (4.5) shows that the estimation error is a deterministic function of time.

Define an innovation process as the normalized deviation of the return from its conditional mean:

$$\begin{aligned} dZ'_t &= \sigma_1^{-1}\left(\frac{dS_t}{S_t} - m_t dt\right), \\ \text{with} \quad Z'_0 &= 0. \end{aligned} \tag{4.6}$$

Gennotte [1986] shows that Z'_t is an observable Brownian motion, and contains the same information as that in the path of S_t. Given the investor's estimation of the expected return, and the introduction of Z'_t, the risky asset as perceived by the investor is given by

$$\frac{dS_t}{S_t} = m_t dt + \sigma_1 dZ'_t, \tag{4.7}$$

$$dm_t = \theta(\bar{\mu} - m_t)dt + \frac{\sigma_1\sigma_2\rho + V_t}{\sigma_1} dZ'_t. \tag{4.8}$$

Now turn to the second step of the investor's decision, the investment optimization problem in which she uses her current estimate of μ_t to choose the optimal plan of consumption and asset allocation. Denote the initial wealth at time 0 of the investor by W_0. Her consumption plan is characterized by an consumption-rate process c_t, and her portfolio plan is a process of portfolio weights in the stock x_t. The residual, $1 - x_t$, is allocated to the riskless bond. From the self-financing property of the portfolio strategy, it follows that the wealth process is given by

$$dW_t = [r + x_t(m_t - r)]W_t dt - c_t dt + x_t\sigma_1 W_t dZ'_t. \tag{4.9}$$

Assume that the investor has constant relative risk aversion. Then the optimization problem faced by the investor is formulated as the following:

$$\begin{aligned} &\sup_{x_t, c_t} E\left[\int_0^T e^{-\eta t}\frac{c_t^{1-\gamma}}{1-\gamma}dt + \psi e^{-\eta T}\frac{W_T^{1-\gamma}}{1-\gamma}\right] \\ \text{s.t.} \quad & dW_t = [r + x_t(m_t - r)]\, W_t dt - c_t dt + x_t\sigma_1 W_t dZ'_t, \\ & W_T \geq 0, \end{aligned} \tag{4.10}$$

where γ is the constant rate of relative risk aversion, η denotes the subjective discount rate, and ψ denotes the importance of bequest to the investor.

Assuming there exists a solution of the dynamic optimization problem (4.10), we can solve the problem. Thus, the investor's decision problem with the unobservable expected return is solved, which is summarized by the following theorem.

Theorem 4.2.1. *For the investor facing the consumption and investment problem in the presence of state uncertainty, the optimal consumption-wealth ratio, c_t^*/W_t^*, and the optimal portfolio plan, x_t^*, are given by*

$$\frac{c_t^*}{W_t^*} = \frac{1}{F(m_t,t)}, \tag{4.11}$$

$$x_t^* = \frac{1}{\gamma}\frac{m_t - r}{\sigma_1^2} + \frac{\sigma_1\sigma_2\rho + V_t}{\sigma_1^2}\frac{1}{F(m_t,t)}\frac{\partial F}{\partial m_t}(m_t,t), \tag{4.12}$$

where $F(m_t,t)$ is the solution of the following partial differential equation (PDE):

$$\begin{aligned} -\frac{\partial F}{\partial t}(m_t,t) &= 1 - \left[\frac{\eta}{\gamma} - \frac{1-\gamma}{\gamma}r - \frac{1-\gamma}{2\gamma^2}\left(\frac{m_t - r}{\sigma_1}\right)^2\right]F(m_t,t) \\ &\quad + \left[\theta(\bar{\mu} - m_t) + \frac{1-\gamma}{\gamma}\frac{m_t - r}{\sigma_1}\frac{\sigma_1\sigma_2\rho + V_t}{\sigma_1}\right]\frac{\partial F}{\partial m_t}(m_t,t) \\ &\quad + \frac{1}{2}\left(\frac{\sigma_1\sigma_2\rho + V_t}{\sigma_1}\right)^2\frac{\partial^2 F}{\partial m_t^2}(m_t,t) \end{aligned} \tag{4.13}$$

with the boundary condition

$$F(m_T,T) = \psi^{1/\gamma}.$$

Proof. See Appendix 4.A. □ □

The function $F(m_t,t)$ gives the value of the optimal wealth-consumption ratio, as is illustrated by (4.11). The first term in (4.12) is the familiar myopic allocation to stock (the allocation when $\gamma = 1$), and the second term is the stock allocation for intertemporal hedging purposes. It is well known that intertemporal hedging demands arise from stochastic investment opportunity sets (Merton [1971]). Because the investment opportunity set in this model is stochastic, and perceived as such by the investor, it is natural to have an intertemporal hedging term in the optimal allocation.

The investor's consumption and investment depends on her assessment of the investment opportunity, m_t. Moreover, Equation (4.8) shows that the investor takes into account the whole history of the risky asset return till now in her assessment of the current investment opportunity, for the reason that Z_t' incorporates the same information as the path of S_t does. Therefore, the consumption and investment of the investor depend on the path of the risky asset return over the observation interval. The path-dependence of the consumption and investment implies that the investor first takes some form of "average" of the historical returns on the risky asset, and then decides her consumption and investment on the basis of the average.

4.3 The optimal strategy in the stable phase

In the presence of state uncertainty, the investor's assessment of the investment opportunity set is generally imperfect in the sense that the estimation error, V_t, cannot be ignored even if realized returns are observed continuously over an infinite horizon. In this case, as the observation interval tends to $[0, \infty)$, the variance of the estimator converges to a constant given by

$$\bar{V} = \sigma_1^2 \left[\sqrt{\theta^2 + \frac{\sigma_2^2}{\sigma_1^2} + 2\theta\rho\frac{\sigma_2}{\sigma_1}} - \left(\theta + \rho\frac{\sigma_2}{\sigma_1} \right) \right]. \tag{4.14}$$

We refer to the phase in which the variance of the estimator is equal to this constant by the *stable phase*. Once in this phase, the investor will continue to update the estimate of the expected return from observing realized returns, m_t, but she cannot improve upon the estimation error, V_t. An interesting feature in the stable phase is that the return on the risky asset is perceived by the investor to be less volatile than if there is no state uncertainty, as is shown in the following proposition.

Proposition 4.3.1. *In the stable phase, the expected return of the risky asset is perceived to have lower volatility than that of the latent expected return μ_t, namely,*

$$\left| \frac{\sigma_1\sigma_2\rho + \bar{V}}{\sigma_1} \right| \leq \sigma_2. \tag{4.15}$$

Proof. From Equation (4.8), it follows that in the stable phase, the estimate of the expected return is given by

$$dm_t = \theta(\bar{\mu} - m_t)dt + \kappa\sigma_1 dZ_t', \tag{4.16}$$

where

$$\kappa := \frac{\sigma_1\sigma_2\rho + \bar{V}}{\sigma_1^2}. \tag{4.17}$$

Compared to the stochastic differential equation (SDE) (4.3) that characterizes the unobservable expected return, this SDE is also of Ornstein-Uhlenbeck type with the same drift term, but a different diffusion term. Thus, a lower diffusion term as presented by (4.15) means that the perceived expected return, m_t, is less volatile than the latent expected return, μ_t. After substituting the value of $\bar{V}$ given by (4.14) into (4.15), the inequality is equivalent to

$$\theta^2 + \frac{\sigma_2^2}{\sigma_1^2} - 2\theta\frac{\sigma_2}{\sigma_1} \leq \theta^2 + \frac{\sigma_2^2}{\sigma_1^2} + 2\theta\rho\frac{\sigma_2}{\sigma_1} \leq \theta^2 + \frac{\sigma_2^2}{\sigma_1^2} + 2\theta\frac{\sigma_2}{\sigma_1},$$

which obviously holds since $|\rho| \leq 1$. $\square$ $\square$

An investigation of the investor's decision in the stable phase is of interest since it leads to insights on how an investor with "sufficient experience" will consume and invest. From the perspective of solving the dynamic asset allocation problem with state uncertainty, the stable phase allows for a closed-form solution when the investor is more risk averse than log utility ($\gamma > 1$). This assumption that $\gamma > 1$ is empirically relevant, since many empirical studies of people's risk aversion (see e.g. Friend and Blume [1975], Pindyck [1988], and Szpiro [1986]) and the literature on the equity premium puzzle lend support to it.

In the stable phase, the PDE in Theorem 4.2.1 is reduced to

$$\begin{aligned} -\frac{\partial F}{\partial t}(m_t, t) = 1 - \left[\frac{\eta}{\gamma} - \frac{1-\gamma}{\gamma}r - \frac{1-\gamma}{2\gamma^2}\left(\frac{m_t - r}{\sigma_1}\right)^2\right] F(m_t, t) \\ + \left[\theta(\bar{\mu} - m_t) + \frac{1-\gamma}{\gamma}\kappa(m_t - r)\right]\frac{\partial F}{\partial m_t}(m_t, t) + \frac{1}{2}\kappa^2\sigma_1^2\frac{\partial^2 F}{\partial m_t^2}(m_t, t), \quad (4.18) \end{aligned}$$

which can be solved in closed form as follows. First, enlightened by the solutions of the PDEs in Kim and Omberg [1996] and Wachter [2002], we take the following as the guessed form of solution of (4.18):

$$F(m_t, t) = \int_t^T e^{H(m_t, s-t)} ds + \psi^{1/\gamma} e^{H(m_t, T-t)}, \quad (4.19)$$

where

$$H(m_t, \tau) = -\frac{\eta}{\gamma} + \frac{1-\gamma}{\gamma}\left[\frac{1}{2}A_1(\tau)\left(\frac{m_t - r}{\sigma_1}\right)^2 + A_2(\tau)\left(\frac{m_t - r}{\sigma_1}\right) + A_3(\tau)\right].$$

The three functions, A_1, A_2, and A_3, can be solved by substituting (4.19) back into (4.18), and then working out the resultant three differential equations in A_1, A_2, and A_3. The assumption that $\gamma > 1$ is needed to ensure the existence of well behaved solution of these differential equations. The method for solving these equations is standard in the literature, and the solutions are:

$$\begin{aligned}
A_1(\tau) &= \frac{1}{\gamma}\frac{2\left(1-e^{-q\tau}\right)}{2q-(q+b_1)\left(1-e^{-q\tau}\right)},\\
A_2(\tau) &= \frac{1}{\gamma}\frac{4\theta(\bar{\mu}-r)\left(1-e^{-q\tau/2}\right)^2}{\sigma_1 q[2q-(q+b_1)\left(1-e^{-q\tau}\right)]},\\
A_3(\tau) &= r\tau+\int_0^\tau\left[\frac{b_2}{2}A_2(x)^2+\frac{\theta(\bar{\mu}-r)}{\sigma_1}A_2(x)+\frac{b_2}{2}\frac{\gamma}{1-\gamma}A_1(x)\right]dx,
\end{aligned}$$

where

$$\begin{aligned}
b_0 &= \frac{1}{\gamma},\\
b_1 &= 2\left(\frac{1-\gamma}{\gamma}\kappa-\theta\right),\\
b_2 &= \frac{1-\gamma}{\gamma}\kappa^2,\\
q &= \sqrt{b_1^2-4b_0b_2}.
\end{aligned}$$

Therefore, the optimal strategy of the investor is solved as is presented in the following lemma.

Lemma 4.3.2. *In the stable phase, the optimal consumption strategy is given by*

$$\frac{c_t^*}{W_t}=\left[\int_t^T e^{H(m_t,s-t)}ds+\psi^{1/\gamma}e^{H(m_t,T-t)}\right]^{-1}. \tag{4.20}$$

The optimal portfolio strategy is given by

$$\begin{aligned}
x_t^* &= \frac{1}{\gamma}\frac{m_t-r}{\sigma_1^2}\\
&\quad+\frac{1-\gamma}{\gamma}\frac{\kappa}{\sigma_1}\frac{\int_t^T[A_1(s-t)\frac{m_t-r}{\sigma_1}+A_2(s-t)]e^{H(m_t,s-t)}ds}{\int_t^T e^{H(m_t,s-t)}ds+\psi^{1/\gamma}e^{H(m_t,T-t)}}\\
&\quad+\frac{1-\gamma}{\gamma}\frac{\kappa}{\sigma_1}\frac{\psi^{1/\gamma}[A_1(T-t)\frac{m_t-r}{\sigma_1}+A_2(T-t)]e^{H(m_t,T-t)}}{\int_t^T e^{H(m_t,s-t)}ds+\psi^{1/\gamma}e^{H(m_t,T-t)}}.
\end{aligned} \tag{4.21}$$

Note that the second and third terms in (4.21) are the allocation for intertemporal hedging purposes. And the third term arises from bequest motives, for this term disappears if the investor does not derive utility from bequest ($\psi = 0$).

While (4.20) and (4.21) may first appear complicated, they can used to provide insights into the decision of the experienced investor in the presence of state uncertainty.

Proposition 4.3.3. *Assume that the investor is more risk averse than log utility ($\gamma > 1$), and that the long-term mean of equity risk premium is positive ($\bar{\mu} - r > 0$). Then as long as the perceived risk premium, $m_t - r$, is positive, the optimal consumption-wealth ratio is increasing in the perceived expected return m_t.*

The proof of this proposition is straightforward. Appendix 5.A shows that $A_1(\tau)$ and $A_2(\tau)$ are both positive when $\gamma > 1$ and $\bar{\mu} - r > 0$. Then as long as $m_t - r > 0$, the partial derivative of $H(m_t, \tau)$ with respect to m_t is negative. In this situation, it immediately follows that the partial derivative of (4.20) with respect to m_t is positive. Therefore the optimal consumption-wealth ratio increases with m_t. This result is also intuitive, in that people tend to consume more when they believe economic perspectives are good.

Proposition 4.3.4. *Assume that the investor's preference and the financial market dynamics satisfy the following conditions:*

1. *the investor is more risk averse than log utility ($\gamma > 1$);*
2. *the long-term mean of equity risk premium is positive ($\bar{\mu} - r > 0$);*
3. *$\kappa < 0$, or equivalently $\sigma_2/\sigma_1 + 2\theta\rho < 0$.*

Then as long as the perceived risk premium, $m_t - r$, is positive, the optimal investment policy has the following properties:

1. *the intertemporal hedging demand is positive, namely, the allocation to the risky asset is higher than the corresponding myopic allocation;*
2. *the optimal allocation increases with the investment horizon.*

Conversely, if $\kappa > 0$, or equivalently $\sigma_2/\sigma_1 + 2\theta\rho > 0$ while other things being equal, then the opposite is true. In the special case where $\kappa = 0$, or $\sigma_2/\sigma_1 + 2\theta\rho = 0$, the optimal investment plan corresponds to the myopic one.

Proof. See Appendix 4.A. □ □

Out of the three conditions in Proposition 4.3.4, the first two are straightforward. The third one, however, is more delicate, and a further exploration is in order. This condition, which concerns the sign of $\sigma_2/\sigma_1 + 2\theta\rho$, determines whether the horizon effect is positive, negative or absent. The conventional wisdom suggests a positive horizon effect by saying that a long-horizon investor should invest more in equity. The reason is that equity is less risky in long horizons as equity risk premium is time-varying and above-average returns tend to offset below-average returns. Along this line of reasoning, the condition that $\sigma_2/\sigma_1 + 2\theta\rho < 0$ should imply equity's reduced risk over long horizon, and vice versa. As is shown in Appendix 4.A, it is indeed this case. Only if $\sigma_2/\sigma_1 + 2\theta\rho < 0$, the stock in the perception of the investor is less risky over long horizons as evidenced by the fact that the stock return's average variance decreases with horizons. Note that in the case where $\sigma_2/\sigma_1 + 2\theta\rho > 0$, the stock is actually perceived more risky when the investor's horizon is longer. In this situation, the risk-return profile of the stock gets worse in long horizons, and it is understandable for the investor to decrease the stock allocation with the horizon, implying a negative horizon effect.

The conventional wisdom of more equity investment for long-term investors suggests that financial market dynamics satisfying $\sigma_2/\sigma_1 + 2\theta\rho < 0$ are more likely. For this condition to hold, it is necessary though not sufficient that the correlation between the stock return and the risk premium is negative ($\rho < 0$). When this correlation is negative, the conventional wisdom of positive horizon effect tends to prevail if the expected return is stable (small σ_2 and large θ) and the stock return is volatile (large σ_1).

Thanks to the presence of the parameter, ψ, which measures the strength of bequest motives, this model allows for an analysis on how people's bequest motives affect their consumption and investment. From (4.20) and the positiveness of γ, we have the following proposition.

Proposition 4.3.5. *The optimal consumption-wealth ratio is decreasing in the strength of bequest motives, ψ.*

This result is what one would expect: the more you want to leave as bequests, you should consume less in order to have more to save as bequests.

Proposition 4.3.6. *Assume that (i)$\gamma > 1$; (ii) $\bar{\mu}-r > 0$; and (iii) $\sigma_2/\sigma_1+2\theta\rho < 0$. Then as long as the perceived risk premium, m_t-r, is positive, the optimal allocation to the risky asset increases with the strength of bequest motives, ψ. Conversely, if $\sigma_2/\sigma_1 + 2\theta\rho > 0$ while other things being equal, then the opposite is true.*

The derivative of the optimal stock allocation x_t^* with respect to ψ is

$$\psi^{\frac{1}{\gamma}-1}\frac{1-\gamma}{\gamma^2}\frac{1}{\sigma_1}\frac{\sigma_1\sigma_2\rho+\bar{V}}{\sigma_1^2}$$
$$\frac{\left[A_1(\tau)\frac{m_t-r}{\sigma_1}+A_2(\tau)\right]\int_0^\tau e^{H(m_t,s)}ds-\int_0^\tau[A_1(s)\frac{m_t-r}{\sigma_1}+A_2(s)]e^{H(m_t,s)}ds}{\left[\int_0^\tau e^{H(m_t,s)}ds+\psi^{1/\gamma}e^{H(m_t,\tau)}\right]^2},$$

where $\tau = T-t$. With the three assumptions in the above proposition, it follows from (4.34) that as long as $m_t - r > 0$, this derivative is positive. And the sign of the derivative is reversed when ceteris paribus, the sign of $\sigma_2/\sigma_1 + 2\theta\rho$ turns from negative to positive. Note that the effect of bequest motives on the stock allocation has the same pattern as that of the investment horizon. In other words, the stock allocation increases with the strength of bequest motives when there is a positive horizon effect, and vice versa. An intuitive explanation goes as follows. Analogous to the notion of duration in fixed income, an investor with a strong bequest motive should have a longer "effective horizon" than an investor who has the same characteristics but derives little utility from bequests as the former distributes more wealth on the fixed horizon date. Thus an increase in bequest motives, in effect, prolongs the horizon, which explains the close link between the strength of bequest motives, ψ, and the investment horizon.

4.4 Calibration and discussion

In this section, we present some representative numerical illustrations and discussions. The parameter values underlying the calibration exercises are drawn from Barberis [2000] and Wachter [2002], and summarized in Table 4.1.

Parameter Description	Notation	Parameter Values
Subjective discount rate	η	0.0624
Riskless rate	r	0.0168
Volatility of stock price	σ_1	0.1510
Volatility of expected return	σ_2	0.0343
Unconditional mean of expected return	$\bar{\mu}$	0.0576
Mean-reverting parameter of expected return	θ	0.2712
Correlation coefficient	ρ	-0.9351

Table 4.1: **The parameter values used in the numerical analysis** The parameter values are calculated based on Barberis [2000] and Wachter [2002]. All parameters are in annual units.

Figure 4.1 shows how the estimation error V_t changes with the length of observation interval for different values of the initial estimation error V_0. Regardless of the various

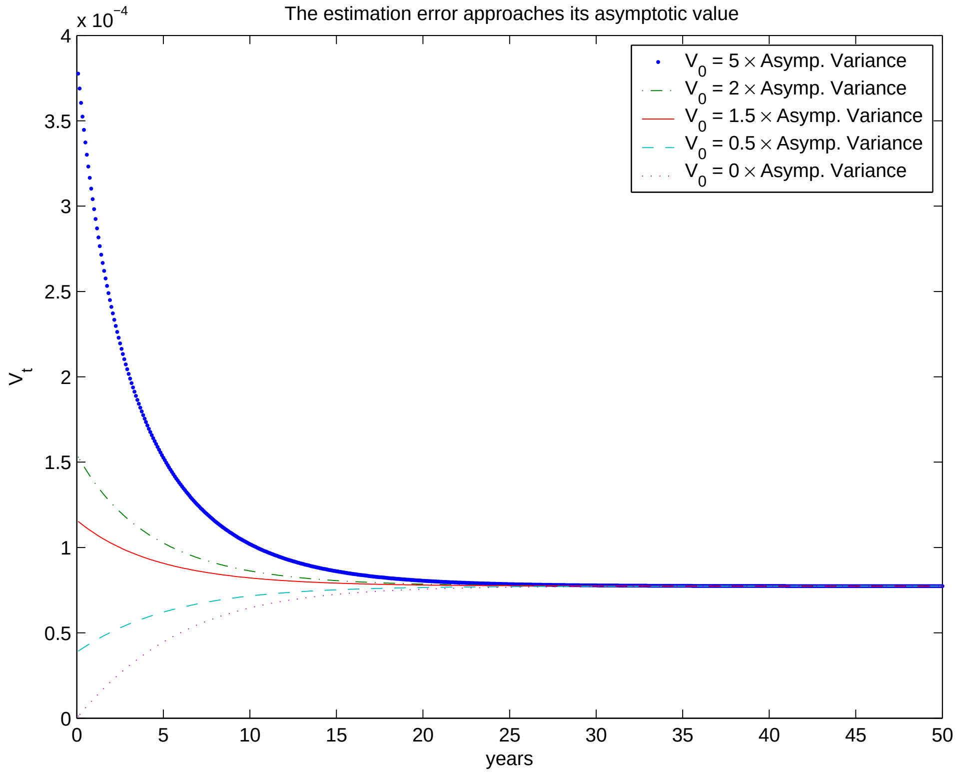

Figure 4.1: **The change of estimation error, V_t, over the observation interval** This figure shows how estimation error V_t changes with the observation interval for different values of the initial estimation error V_0: $V_0 = a\bar{V}$, where $a = \{5, 2, 1.5, 0.5, 0\}$. The variance is in annual units, and the parameter values are given in Table 4.1.

values of V_0 considered here, it takes about 30 years for the estimation error to come close to its asymptotic value. Thus in this numerical setting, the investor is able to reach the stable phase with a fairly short history of the risky asset returns. For this reason, we shall devote the numerical analysis to the stable phase.

To gain some insight on the magnitude of the difference between the estimated and latent expected returns discussed in Proposition 4.3.1, we compare these two returns in this numerical setting. From (4.30) it follows that the estimated expected return in the stable phase is asymptotically stationary and normal: $\Phi\left(\bar{\mu}, \frac{\kappa^2\sigma_1^2}{2\theta}\right)$. Because the parameters characterizing the dynamics of the latent expected return in (4.3) are assumed to be known to the investor, she knows that the asymptotic distribution of μ_t is also normal with $\Phi\left(\bar{\mu}, \frac{\sigma_2^2}{2\theta}\right)$. Figure 4.2 shows that in this numerical setting, the estimated expected return is appreciably less volatile than the latent expected return. For instance, the probability of negative latent risk premium is about 0.19 while it is 0.12 for the estimated risk premium. The investor is aware of the difference between the unconditional

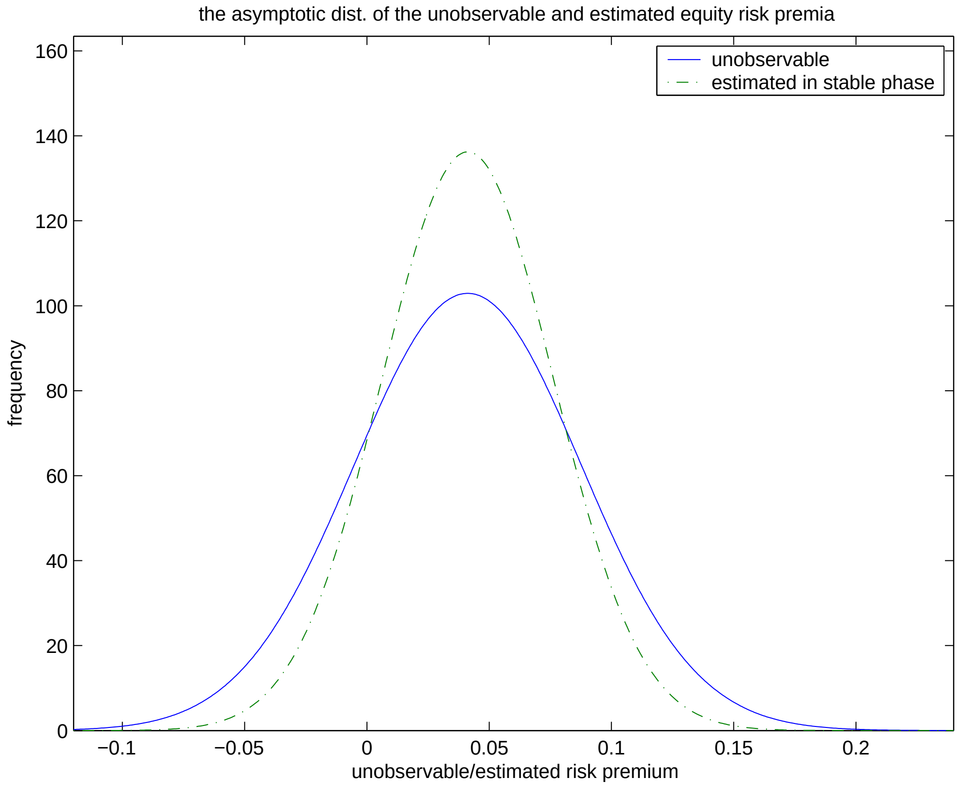

Figure 4.2: **The asymptotic distributions of the unobservable risk premium and the estimated risk premium in the stable phase** This figure shows the asymptotic distributions of the unobservable risk premium and the estimated risk premium in the stable phase. The risk premia are in annual rates, and the parameter values are given in Table 4.1.

distributions of the estimated and latent expected returns though, she, in the face of state uncertainty, seems to opt to be prudent in estimation.

4.4.1 Optimal investment and consumption policy

Figure 4.3 shows the optimal consumption and investment plan for a range of estimated values of expected return. The value range corresponds to the 95% confidence interval with respect to the asymptotic distribution of m_t. Observe first that the more risk tolerant the investor is (lower γ), the more responsive the consumption and investment are to the variation of the estimated return. Secondly, consistent with Proposition 4.3.3, the consumption is increasing with the estimated return. Likewise, a higher estimated return leads the investor to allocate more to the stock.

Recall that the investor forms the estimate of the expected return on the basis of the realized returns on the stock, and the current estimate depends on the entire history of the realized returns. Thus, any shocks to the realized returns in the past will affect

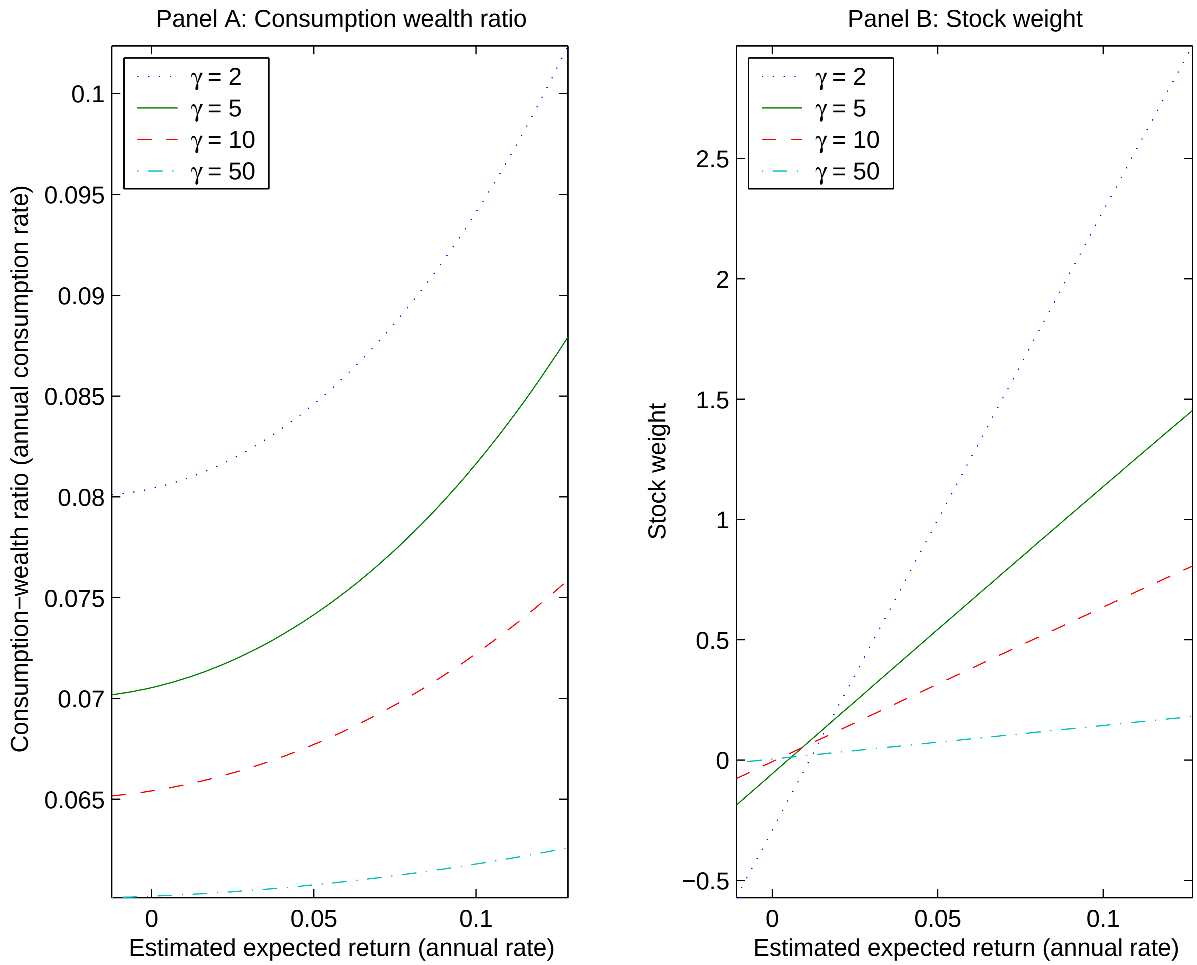

Figure 4.3: **How the optimal consumption and investment respond to the estimated expected return** This figure shows how the optimal consumption-wealth ratio (Panel A) and allocation to the risky asset (Panel B) change with the estimated expected return for different rates of relative risk aversion ($\gamma = \{2, 5, 10, 50\}$). It is assumed that the horizon T is 20 years, and $\psi = 1$. Other parameter values are given in Table 4.1.

the investor's current estimate. It's intuitively clear that other things being equal, a big shock to the stock price will result in a relatively more significant adjustment of the estimate. Moreover, given the size of the shock, the current estimate also depends on how long ago the shock happened, or the "age" of the shock. As is illustrated by Figure 4.4, the impact of the shock on the current estimate declines exponentially with its age. The impact on the estimate in this figure is computed as follows. First we compute the estimated expected for a baseline path of the realized returns dS_t/S_t. Then introduce a shock of some size to a certain time point of this path. We vary the time of the shock, get the corresponding new estimate, and compute the difference between the new estimate and the baseline estimate, which measures the impact of the stock price shock of varying ages.

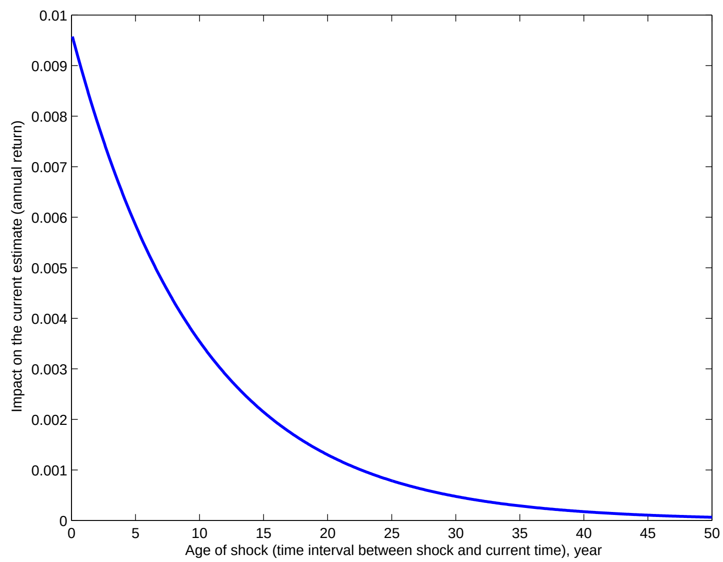

Figure 4.4: **How the age of the shock to the stock price affects the estimate of the expected return** This figure shows how the age of the shock to the stock price affects the estimated of expected return. The size of the shock is assumed to be -1.2816 realization of $Z_{S,t+1} - Z_{S,t}$ (recall that 1.2816 is the 90 percentile of a standard normal variable), and the age of the shock varies from 1 month to 50 years. The impact on the estimated expected return is represented in yearly units. Other relevant parameter values are given in Table 4.1.

The exponential relationship between the age of the shock and the impact of the shock on the estimate illustrated in Figure 4.4 can also be established analytically. From Equation (4.4) and $V_t = \bar{V}$ in the stable phase, it follows that

$$m_t = \bar{\mu} + e^{-(\theta+\kappa)t}(m_0 - \bar{\mu}) + \kappa \int_0^t e^{-(\theta+\kappa)(t-u)} \left(\frac{dS_u}{S_u} - \bar{\mu} du \right).$$

It is clear from the above equation that the impact of the stock price shock, dS_u/S_u, on the estimate declines with the age of the shock, $t - u$.

In this numerical exercise, we introduce a negative shock to the stock price, which leads the investor to *increase* the estimate of expected return for the reason that κ is negative in this setting (cf. (4.6) and (4.30)). As Figure 4.3 shows, an increase in the estimated expected return will typically result in higher consumption-wealth ratio and more allocation to the risky asset. Integrating the findings from Figure 4.4 and 4.3, we can expect that the age of the stock price shock also has an impact on the investor's consumption and investment. To illustrate the magnitude of this impact, Figure 4.5 presents the percentage changes in the consumption-wealth ratio and the stock weight caused by the shock of varying ages. Observe that a recent price shock has a appreciable impact on the investor's consumption and investment, and the impact is decreasing noticeably with the age of the shock.

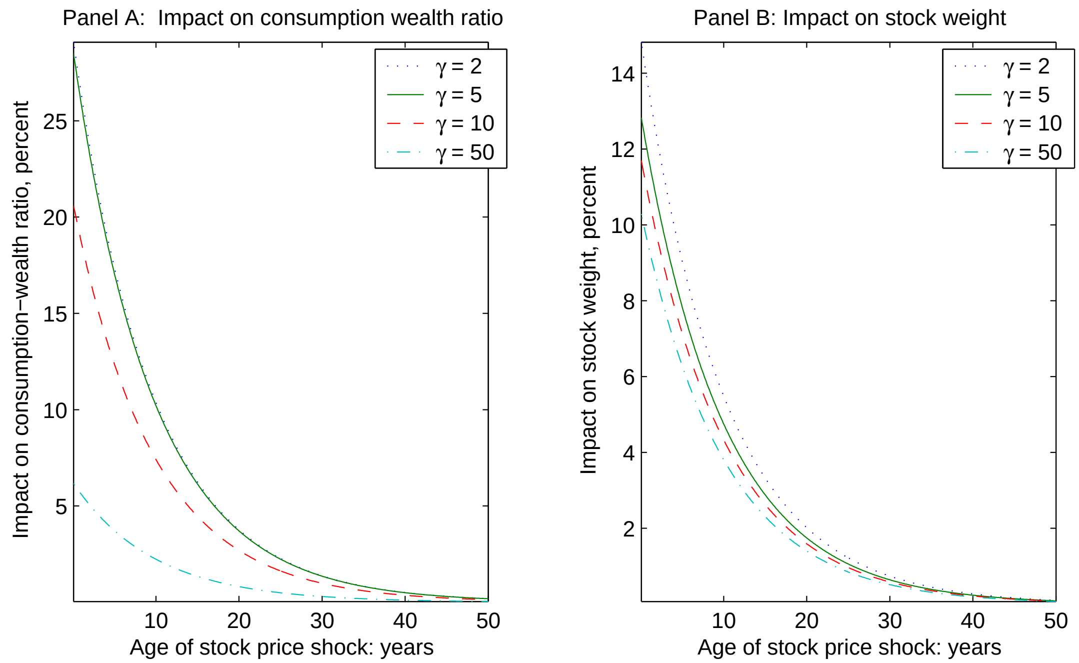

Figure 4.5: **How the age of the shock to the stock price affects the optimal consumption and investment** This figure is based on a baseline path of constant stock returns equal to $\bar{\mu}$. The size of the shock is assumed to be -1.2816 realization of $Z_{S,t+1} - Z_{S,t}$ (recall that 1.2816 is the 90 percentile of a standard normal variable), and the age of the shock varies from 1 month to 50 years. The impact on the consumption and investment is expressed as the percentage changes in consumption-wealth ratio and stock weight with respect to their baseline values. It is assumed that the horizon is 20 years and $\psi = 1$. Other relevant parameter values are given in Table 4.1.

Turn to the investor's stock allocation for intertemporal hedging purposes. We assume that the current estimated expected return is equal to the long-term mean $\bar{\mu}$. As Proposition 4.3.4 predicts, with a negative value of $\sigma_2/\sigma_1 + 2\rho\theta$ the hedging allocation is positive and increasing in the horizon (Figure 4.6). This figure also shows that the hedging allocation is not monotonic in the rate of relative risk aversion.

4.4.2 Implications for the benefit policy of pension funds

Apart from portfolio management, a collective defined-contribution (CDC) pension fund also has the task of deciding the benefit policy, especially the amount of benefit paid out in a certain situation. When a CDC fund is modeled in the framework of Merton's problem like the one discussed in this chapter, the consumption policy in Merton's problem can be interpreted as the benefit policy of the pension fund. In the following, we consider the implications of this study for the benefit policy of pension funds.

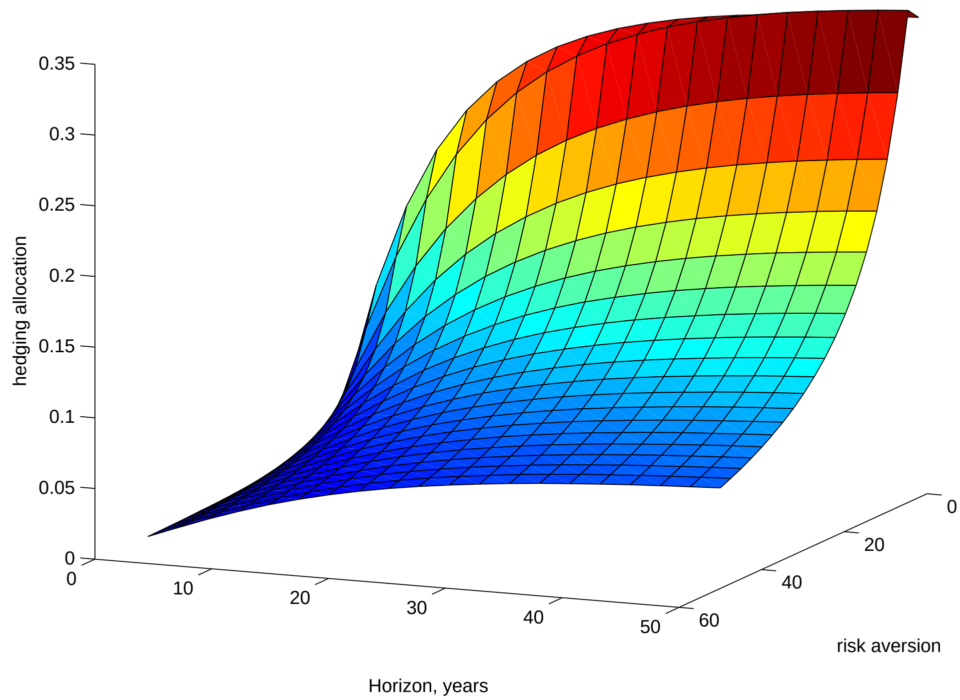

Figure 4.6: **The intertemporal hedging allocation as a function of horizon and risk aversion** This figure shows how the intertemporal hedging allocation changes with the rate of relative risk aversion γ and the investment horizon T. It is assumed that the current estimated expected return equals the long-term mean $\bar{\mu}$, and $\psi = 1$. Other relevant parameter values are given in Table 4.1.

In this economy with time-varying expected return and state uncertainty, the consumption rate resulting from a given investment and consumption policy can be characterized as[2]

$$\begin{aligned} \frac{dc_t}{c_t} &= \mu_c(m_t)dt + \sigma_c(m_t)dZ_t' \\ &= \mu_c(m_t)dt + \sigma_c(m_t)\sigma_1^{-1}\left(\frac{dS_t}{S_t} - m_t dt\right), \end{aligned} \tag{4.22}$$

where $\mu_c(\cdot)$ and $\sigma_c(\cdot)$ are functions to be specified on the basis of the consumption and investment policy. Observe that $\mu_c(\cdot)$ is the trend of the consumption, whereas $\sigma_c(\cdot)$ represents how the consumption/benefit responds to stock price shocks.

First consider the dynamics of the pension benefit resulting from the optimal policy in the stable phase as given in Lemma 4.3.2. Figure 4.7 shows how the pension benefit/consumption responds to stock price shocks. Observe first that the diffusion term of

[2]In general the function $\mu_c(\cdot)$ also depends on time t. Given the very long horizon typically faced by pension funds, however, we simplify the analysis by assuming that the function $\mu_c(\cdot)$ depends only on the state variable m_t.

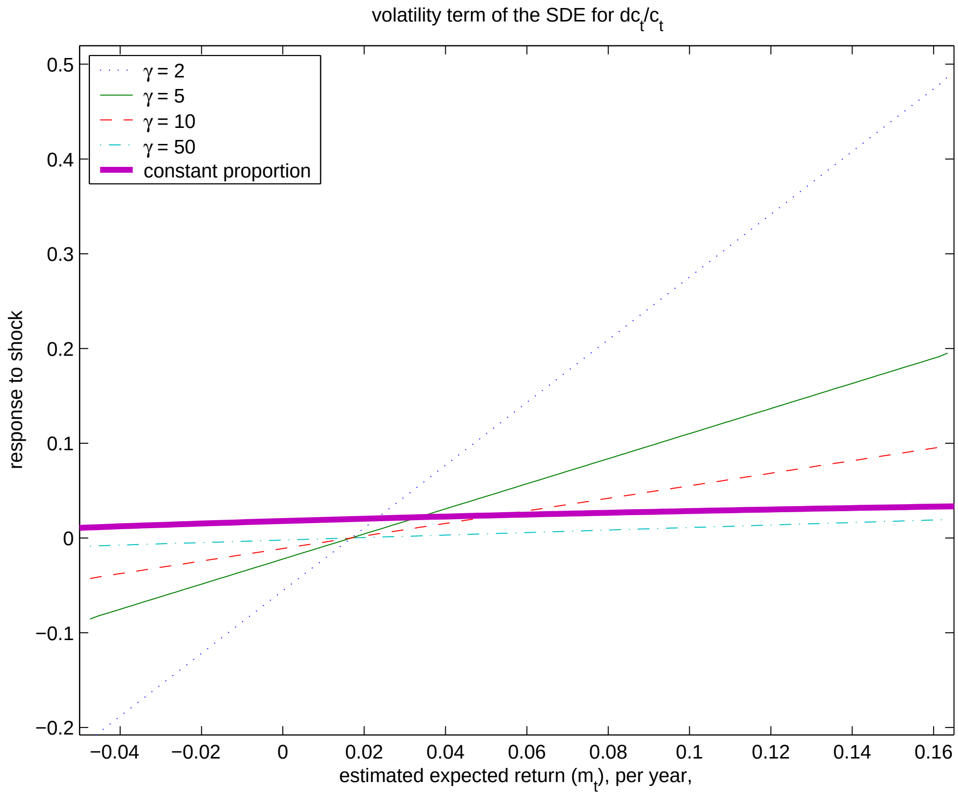

Figure 4.7: **How the pension benefit/consumption responds to stock price shocks** This figure shows the diffusion term $\sigma_c(m_t)$ resulting from different consumption and investment policies. The four light lines are based on the optimal policy given in Lemma 4.3.2 assuming different values of risk aversion ($\gamma \in \{2, 5, 10, 50\}$). The bold solid line assumes that the preference is of Epstein-Zin type with relative risk aversion equal to 5, and the elasticity of intertemporal substitution equal to 1/10, that the investment strategy is a constant proportion strategy with stock weight equal to 50%, and that the agent chooses the optimal consumption plan to maximize the utility. It is assumed that $\psi = 1$. Other relevant parameter values are given in Table 4.1.

the SDE, $\sigma_c(m_t)$, is increasing in the estimated expected return m_t regardless of the rates of risk aversion considered here, and that $\sigma_c(m_t)$ is positive when the estimated expected return is large, and negative when the estimated expected return is low. Secondly, the absolute value of $\sigma_c(m_t)$ is increasing with the rate of risk aversion, reflecting that less risk averse investors have relatively more volatile consumption stream.

The figure seems to suggest a linear relationship between m_t and $\sigma_c(m_t)$, which in fact can be established analytically. Applying Ito's formula to (4.20) and (4.30) leads to

$$\sigma_c(m_t) = \left[x_t^* - \kappa \frac{1}{F(m_t,t)} \frac{\partial F}{\partial m_t}(m_t,t) \right] \sigma_1. \tag{4.23}$$

It follows from (4.12) that

$$x_t^* - \kappa \frac{1}{F(m_t,t)} \frac{\partial F}{\partial m_t}(m_t,t) = \frac{1}{\gamma} \frac{m_t - r}{\sigma_1^2}.$$

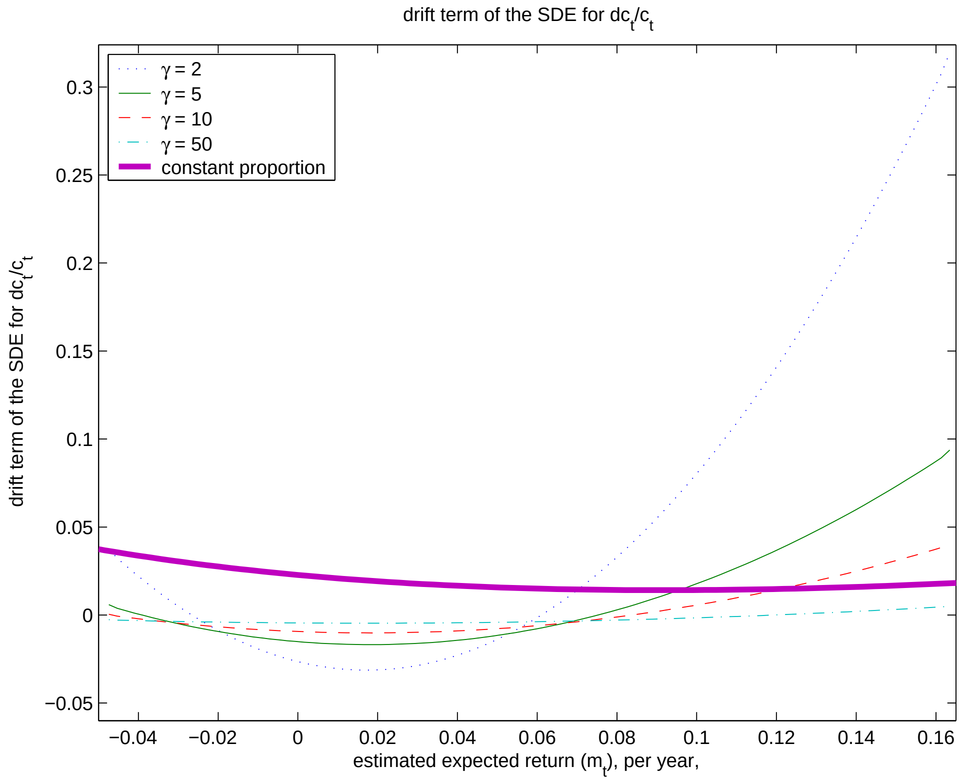

Figure 4.8: **The trend of pension benefit** This figure shows the drift term $\mu_t(m_t)$ resulting from different consumption and investment policy. The four light lines are based on the optimal policy given in Lemma 4.3.2 assuming different rates of risk aversion ($\gamma \in \{2, 5, 10, 50\}$). The bold solid line assumes that the preference is of Epstein-Zin type with relative risk aversion equal to 5, and the elasticity of intertemporal substitution equal to 1/10, that the investment strategy is a constant proportion strategy with stock weight equal to 50%, and that the agent chooses the optimal consumption plan to maximize the utility. It is assumed that $\psi = 1$. Other relevant parameter values are given in Table 4.1.

Recall that the right term in the above equation is the myopic allocation in the stable phase. Thus,

$$\sigma_c(m_t) = \frac{1}{\gamma}\frac{m_t - r}{\sigma_1}. \tag{4.24}$$

Therefore, how the pension benefit/consumption responds to stock price shocks is decided by the rate of risk aversion and the perceived market price of risk, which is closely related to the myopic allocation to the stock.

Figure 4.8 shows the trend of the pension benefit arising from the optimal consumption and investment strategy in the stable phase. Generally, the relationship between the estimated expected return m_t and the drift term $\mu_c(m_t)$ exhibits a "U" shape. The trend of the pension benefit is the lowest when the estimated expected return is around the riskless interest rate. The trend increases when the estimated expected either increases to significantly positive region or drops to the negative region. It is intuitively plausible for the trend of the pension benefit to increase when the estimated expected return drops in the negative region, for the reason that the agent is allowed to short the stock to take advantage of negative expected return.

Combining the findings from Figure 4.7 and 4.8, we can see that whenever the trend of the pension benefit is low, the benefit is relatively less responsive to stock price shocks, and thus less volatile. Conversely, when the trend is high, the benefit is relatively more responsive and then more risky. This pattern reflects that the agent trades off the amount and riskiness of the consumption stream.

4.4.3 Fixed-mix policies

Of course, the investment strategies that have been developed above depend rather strongly on model assumptions. Given the uncertainty that surrounds such assumptions, pension funds may consider it wise not to try to time the market and may prefer to hold on to a policy in which for instance a fixed percentage of wealth is invested in risky assets. When the investment policy is fixed in this way, there is still the consumption policy to be determined. Fund policy should aim for smooth consumption patterns; however, under a fixed-mix policy which involves a substantial investment in risky assets, gradual adjustment to varying market circumstances is necessary. Taking a long-horizon perspective, we may assume that consumption policy does not depend on calendar time. More specifically, suppose that consumption at time t is given by

$$c_t = \phi(m_t) W_t \tag{4.25}$$

where $\phi(\cdot)$ is a differentiable policy function. It is then readily verified that, under the model of Section 4.2,

$$\frac{dc_t}{c_t} = \frac{\phi'(m_t)}{\phi(m_t)} dm_t + \kappa x \sigma_1^2 \, dt + \frac{dW_t}{W_t}$$

where x denotes the fraction of wealth invested in risky assets. Therefore the relative (instantaneous) volatility of consumption is given by

$$\sigma_c = \Big(x + \frac{\phi'(m_t)}{\phi(m_t)} \kappa \Big) \sigma_1. \tag{4.26}$$

Typically, both x and $\phi'(m)$ are positive; a reduction of volatility will then occur if κ is negative and not too large in absolute value.

To find specific values, a numerical optimization was carried out using Epstein-Zin preferences given recursively by

$$V_t = \Big(\eta\Delta t\, c_t^{1-1/\psi} + (1-\eta\Delta t)\big(E_t(V_{t+\Delta t}^{1-\gamma})\big)^{\frac{1-1/\psi}{1-\gamma}}\Big)^{\frac{1}{1-1/\psi}}$$

with relative risk aversion parameter $\gamma = 5$, elasticity of intertemporal substitution $\psi = 0.05$, and time preference parameter $\eta = 0.06$; the length of the horizon is taken long enough so that convergence to stationary policies takes place. The optimal policy depends on the fraction of wealth held in risky assets. It appears that in each case the optimal consumption/wealth ratio is close to being linear in the business cycle variable m_t, so that we consider policy functions of the form

$$\phi(m_t) = \alpha + \beta(m - \bar{\mu})/\sigma_1. \tag{4.27}$$

For the case in which 50% of wealth is held in risky assets, we find $\alpha \approx 0.035$ and $\beta \approx 0.0075$. Under a more conservative asset mix, with 20% of wealth held in risky assets, we obtain $\alpha \approx 0.025$ and $\beta \approx 0.0025$. Figure 4.9 shows the results of simulations with these policies. It is seen that the 50/50 asset mix leads to a larger spread of consumption, but that it improves in many cases on consumption under the 20/80 asset mix. On the basis of (4.26), it is found that in both cases the relative volatility of consumption is about half of what it would be if the consumption-wealth ratio would be constant.

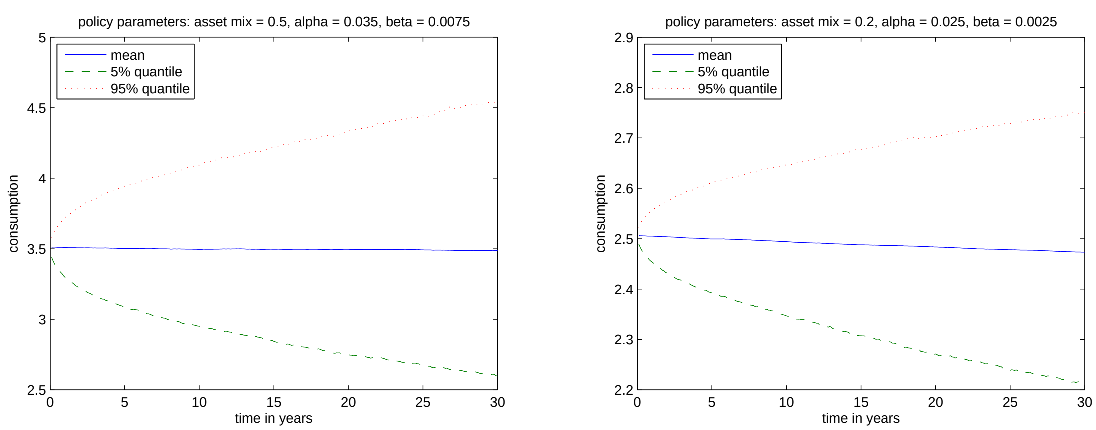

Figure 4.9: **Consumption** The figures show the average behavior in time of consumption under the rule (4.27), as well as 5% and 95% quantiles.

The response to stock price shocks and the trend of the consumption are shown in Figures 4.7 and 4.8. Slightly increasing with the estimated expected return, the term $\sigma_c(m_t)$ is relatively invariable in different states of the economy (Figure 4.7). An intuitive explanation is that the exposure to risk is stable over different states of the economy due to the constant-proportion investment strategy. Observe that in most cases, the absolute value of $\sigma_c(m_t)$ assuming a constant proportion strategy is lower than that arising from the strategy given in Lemma 4.3.2.

4.5 Conclusions and future work

We have analyzed the optimal consumption and investment problem in a setting with time-varying equity premium and state uncertainty, two typical situations faced by investors. This study allows us to investigate their combined implications for consumption and investment. In the spirit of previous work on state uncertainty, the dynamic optimization problem has been addressed in two separate steps: estimation and optimization. The optimal consumption and investment plans have been derived; in particular, the solution can be expressed in closed form in the phase where the investor has a sufficiently long history of the stock price that she can no longer improve upon the estimation error.

The closed-form solution provides insights into the consumption and portfolio choice. For instance, the estimated equity premium is perceived to be less volatile than the unobservable equity premium. The optimal stock allocation may increase, decrease or not vary with the horizon, dependent on the parameter values of the model. From the closed-form solution, it also follows that the optimal consumption-wealth ratio is increasing with the estimated expected return as long as the estimated risk premium is positive. We examine how bequest motives affect the optimal consumption and investment. The consumption is decreasing in the strength of bequest motives. The dependence of stock allocation on bequest motives has the same pattern as the horizon effect.

In the numerical exercise, the investor times the market aggressively in the consumption and investment, and the optimal stock allocation increases with the horizon. We also discuss the implications of this study for the benefit policy of pension funds, and find that the benefit should be relatively stable over time when a constant-proportion investment strategy is employed and the elasticity of intertemporal substitution is low.

This study is only an initial attempt to assess the financial importance of state uncertainty and time-variant equity premia. While we have considered the uncertainty of the state of the economy through the assumption of unobservability of the expected return, we have assumed that the investor knows everything else for sure, including the dynamics of the expected return and all parameters. A more realistic model should incorporate parameter uncertainty. In contrast to studies assuming no state uncertainty, this study examines the case of partial observability. In this study, the *degree* of state uncertainty is kept fixed, but in practice it may be time-varying. A challenging future task is to model the time-variation in the degree of state uncertainty, and investigate its implications for consumption and portfolio decisions.

4.A Appendix

Proof of Theorem 4.2.1

We opt to solve the problem through dynamic programming. The indirect utility function in the intermediate consumption case is

$$J(W_t, m_t, t) = \sup_{(c_s, x_s)_{s\in[t,T]}} E_{W_t, m_t, t}\left[\int_t^T e^{-\eta(s-t)}\frac{c_s^{1-\gamma}}{1-\gamma}ds + \psi e^{-\eta(T-t)}\frac{W_T^{1-\gamma}}{1-\gamma}\right]$$

The Hamilton-Jacobi-Bellman (HJB) equation associated with this dynamic programming problem is

$$\begin{aligned}
\eta J(W_t, m_t, t) \;=\; & \sup_{c_t\geq 0, x_t\in\mathbb{R}}\left\{\frac{c_t^{1-\gamma}}{1-\gamma} + J_t(W_t, m_t, t)\right. \\
& + J_W(W_t, m_t, t)\left[W_t\Big(r + x_t(m_t - r)\Big) - c_t\right] \\
& + \tfrac{1}{2}J_{WW}(W_t, m_t, t)W^2\sigma_1^2 x_t^2 + J_m(W_t, m_t, t)(\bar{\mu} - \theta m_t) \qquad (4.28) \\
& + \tfrac{1}{2}J_{mm}(W_t, m_t, t)\left(\frac{\sigma_1\sigma_2\rho + V_t}{\sigma_1}\right)^2 \\
& \left. + J_{Wm}(W_t, m_t, t)W_t x_t \sigma_1\frac{\sigma_1\sigma_2\rho + V_t}{\sigma_1}\right\},
\end{aligned}$$

with the terminal condition $J(W_T, m_T, T) = \psi\frac{W_T^{1-\gamma}}{1-\gamma}$, where the subscripts of J denote the obvious partial derivatives.

The first-order condition with respect to c_t is

$$c_t^{*-\gamma} = J_W(W_t, m_t, t).$$

Then the optimal consumption strategy is

$$c_t^* = J_W(W_t, m_t, t)^{-\frac{1}{\gamma}}.$$

Similarly, it follows from the first-order condition with respect to x_t that the optimal portfolio strategy is given by

$$x_t^* = -\frac{J_W(W_t, m_t, t)}{W J_{WW}(W_t, m_t, t)\sigma_1^2}(m_t - r) - \frac{J_{Wm}(W, m_t, t)}{W J_{WW}(W, m_t, t)\sigma_1^2}\left[\sigma_1\sigma_2\rho + V_t\right]$$

It's reasonable to guess that the indirect utility function is of the form

$$J(W, m_t, t) = \frac{W^{1-\gamma}}{1-\gamma}F(m_t, t)^\gamma,$$

where

$$F(m_t, t)^\gamma = (1-\gamma)J(1, m_t, t).$$

From the terminal condition $J(W_T, m_T, T) = \psi\frac{W^{1-\gamma}}{1-\gamma}$, it follows that

$$F(m_T, T) = \psi^{1/\gamma}$$

From this guessed form of the indirect utility function, J, and its relevant derivatives, it follows that the optimal consumption-wealth ratio and the optimal portfolio plan are given by (4.11) and (4.12). Finally, inserting the relevant derivatives and the candidate optimal strategy into the HJB equation (4.28) yields the PDE (4.13). □

The explicit solution of V_t and its value in the stable phase

The ODE characterizing V_t (4.5) is of Riccati type, and can be rewritten as

$$dV_t = \left(a_0 + a_1 V_t + a_2 V_t^2\right) dt,$$

where

$$\begin{aligned} a_0 &= \sigma_2^2(1-\rho^2) \\ a_1 &= -2\left(\theta + \rho\frac{\sigma_2}{\sigma_1}\right) \\ a_2 &= -\frac{1}{\sigma_1^2} \end{aligned}$$

This Riccati equation can be solved by recasting it as the following integral form:

$$\int_0^t \frac{1}{a_0 + a_1 V_x + a_2 V_x^2} dV_x = t$$

Generally, the integral depends on the sign of the determinant $4a_0a_2 - a_1^2$. In this case, the determinant is negative. Thus

$$t = -\frac{1}{q}\left[\ln\frac{p + a_1 + 2a_2V_t}{p - a_1 - 2a_2V_t} - \ln\frac{p + a_1 + 2a_2V_0}{p - a_1 - 2a_2V_0}\right]$$

where $p := \sqrt{a_1^2 - 4a_0a_2}$. Therefore,

$$V_t = \frac{e^{-pt}(p + a_1 + 2a_2V_0)(p - a_1) - (p - a_1 - 2a_2V_0)(p + a_1)}{2a_2(p - a_1 - 2a_2V_0) + 2a_2(p + a_1 + 2a_2V_0)e^{-pt}}. \tag{4.29}$$

The value of $\bar{V}$ as given by (4.14) follows immediately from its definition.

Proof of Proposition 4.3.1

From Equation (4.8), it follows that in the stable phase, the estimate of the expected return is given by

$$dm_t = \theta(\bar{\mu} - m_t)dt + \frac{\sigma_1\sigma_2\rho + \bar{V}}{\sigma_1}dZ_t'. \tag{4.30}$$

Compared to the stochastic differential equation (SDE) (4.3) that characterizes the unobservable expected return, this SDE is also of Ornstein-Uhlenbeck type with the same drift term, but a different diffusion term. Thus, a lower diffusion term as presented by (4.15) means that the perceived expected return, m_t, is less volatile than the latent expected return, μ_t. After substituting the value of $\bar{V}$ given by (4.14) into (4.15), the inequality is equivalent to

$$\sqrt{\theta^2 + \frac{\sigma_2^2}{\sigma_1^2} + 2\theta\rho\frac{\sigma_2}{\sigma_1}} \leq \sqrt{\theta^2 + \frac{\sigma_2^2}{\sigma_1^2} + 2\theta\frac{\sigma_2}{\sigma_1}},$$

which obviously holds since $\rho \leq 1$. □

Properties of $A_1(\tau)$ and $A_2(\tau)$

This appendix establishes some properties of $A_1(\tau)$ and $A_2(\tau)$ under the condition that (i) $\gamma > 1$, and (ii) $\bar{\mu} - r > 0$. First define

$$B(\tau) := 2q - (q + b_1)\left(1 - e^{-q\tau}\right).$$

It follows from $\gamma > 1$ that $q > |b_1|$. Thus, $B(\tau) > 0$. Then it follows immediately that

$$A_1(\tau) > 0.$$

With the condition that $\bar{\mu} - r > 0$, one can see that

$$A_2(\tau) > 0.$$

The derivatives of $A_1(\tau)$ and $A_2(\tau)$ are

$$\begin{aligned} A_1'(\tau) &= \frac{1}{\gamma}\frac{4q^2e^{-q\tau}}{B(\tau)^2} \\ A_2'(\tau) &= \frac{1}{\gamma}\frac{4\theta(\bar{\mu} - r)\left(e^{-q\tau/2} - e^{-q\tau}\right)\left[q - b_1 + (q + b_1)e^{-q\tau/2}\right]}{B(\tau)^2} \end{aligned} \tag{4.31}$$

So $A_1'(\tau)$ and $A_2'(\tau)$ are continuous, and

$$\begin{aligned} A_1'(\tau) &> 0, \\ A_2'(\tau) &> 0. \end{aligned}$$

Proof of Proposition 4.3.4

First note that the sign of $\sigma_1\sigma_2\rho + \bar{V}$ is the same as that of $\sigma_2/\sigma_1 + 2\theta\rho$, because

$$\sigma_1\sigma_2\rho + \bar{V} = \sigma_1^2\left[\sqrt{\theta^2 + \frac{\sigma_2}{\sigma_1}\left(\frac{\sigma_2}{\sigma_1} + 2\theta\rho\right)} - \theta\right].$$

Then it follows from the condition that $\sigma_2/\sigma_1 + 2\theta\rho < 0$ that

$$\sigma_1\sigma_2\rho + \bar{V} < 0. \tag{4.32}$$

Assuming $\sigma_2/\sigma_1 + 2\theta\rho < 0$, the hedging term as presented by the second and third term in (4.21) is positive as long as $m_t - r$ is positive, which proves Property 1 in the proposition.

The proof of Property 2 is more involved. First consider the case where $\psi = 0$. In this situation, the optimal strategy in the stable phase (4.21) can be rewritten as

$$x_t^* = \frac{1}{\gamma}\frac{m_t - r}{\sigma_1^2} + \frac{1-\gamma}{\gamma}\frac{\kappa}{\sigma_1} \left[\frac{\int_t^T A_2(s-t)e^{H(m_t,s-t)}ds}{\int_t^T e^{H(m_t,s-t)}ds} + \frac{m_t - r}{\sigma_1}\frac{\int_t^T A_1(s-t)e^{H(m_t,s-t)}ds}{\int_t^T e^{H(m_t,s-t)}ds}\right]$$

Then the derivative of

$$\frac{\int_t^T A_i(s-t)e^{H(m_t,s-t)}ds}{\int_t^T e^{H(m_t,s-t)}ds}, \quad \text{for } i = 1, 2$$

with respect to the investment horizon ($\tau := T - t$) is equal to

$$\frac{e^{H(m_t,\tau)}\left[A_i(\tau)\int_0^\tau e^{H(m_t,s)}ds - \int_0^\tau A_i(s)e^{H(m_t,s)}ds\right]}{\left[\int_0^\tau e^{H(m_t,s)}ds\right]^2} \tag{4.33}$$

Note that because $A_1(\tau)$ and $A_2(\tau)$ are monotonically increasing, it holds that

$$A_i(\tau) > \frac{\int_0^\tau A_i(s)e^{H(m_t,s)}ds}{\int_0^\tau e^{H(m_t,s)}ds}, \quad i = 1, 2. \tag{4.34}$$

Then it follows that the derivative (4.33) is positive, and hence Property 2 expressing a positive horizon effect. In the general case where $\psi \neq 0$, Property 2 can be proved in a similar (but more tedious) way.

Finally, if $\sigma_2/\sigma_1 + 2\theta\rho > 0$, then the argument for these two properties assuming $\sigma_2/\sigma_1 + 2\theta\rho < 0$ is reversed. Therefore, the opposite is true assuming $\sigma_2/\sigma_1 + 2\theta\rho > 0$. □

The riskiness of the stock over long horizons: implications of the sign of $\sigma_1/\sigma_2 + 2\theta\rho$

In the stable phase, the return on the risky asset is characterized by the following linear SDE:

$$d\begin{bmatrix} \ln S_t \\ m_t \end{bmatrix} = \left(\begin{bmatrix} 0 & 1 \\ 0 & -\theta \end{bmatrix}\begin{bmatrix} \ln S_t \\ m_t \end{bmatrix} + \begin{bmatrix} -\frac{1}{2}\sigma_1^2 \\ \theta\bar{\mu} \end{bmatrix}\right) dt + \begin{bmatrix} \sigma_1 \\ \bar{\sigma}_m \end{bmatrix} dZ_t', \tag{4.35}$$

where $\bar{\sigma}_m = \frac{\sigma_1\sigma_2\rho+\bar{V}}{\sigma_1}$. Denote the variance of $\ln S_t$ and m_t at time t by $V_s(t)$ and $V_m(t)$, and the covariance between m_t and $\ln S_t$ by $V_{ms}(t)$. From the linearity of (4.35), it follows that

$$\begin{aligned} \frac{dV_m}{dt}(t) &= \bar{\sigma}_m^2 - 2\theta V_m(t), \\ \frac{dV_{ms}}{dt}(t) &= \sigma_1\bar{\sigma}_m + V_m(t) - \theta V_{ms}(t), \\ \frac{dV_s}{dt}(t) &= \sigma_1^2 + 2V_{ms}(t), \end{aligned} \tag{4.36}$$

with the boundary condition that $V_m(0) = V_{ms}(0) = V_s(0) = 0$. We can solve the system of differential equations. For instance,

$$V_{ms}(t) = \frac{\sigma_1\bar{\sigma}_m}{\theta}\left(1 - e^{-\theta t}\right) + \frac{\bar{\sigma}_m^2}{2\theta^2}\left(1 - e^{-\theta t}\right)^2 \tag{4.37}$$

To characterize how the riskiness of the stock changes with the investment horizons, we define the average variance over the horizon τ by

$$AV(\tau) = \frac{V_s(\tau)}{\tau}.$$

Then

$$\begin{aligned} \frac{dAV}{d\tau}(\tau) &= \frac{1}{\tau}\left[\sigma_1^2 + 2V_{ms}(t)\right] - \frac{1}{\tau^2}\int_0^\tau \left[\sigma_1^2 + 2V_{ms}(x)\right] dx \\ &= \frac{1}{\tau}\left[\sigma_1 + \frac{\bar{\sigma}_m}{\theta}\left(1 - e^{-\theta t}\right)\right]^2 - \frac{1}{\tau^2}\int_0^\tau \left[\sigma_1 + \frac{\bar{\sigma}_m}{\theta}\left(1 - e^{-\theta x}\right)\right]^2 dx, \end{aligned}$$

where the second equality follows from (4.37). Appendix 4.A shows that $\bar{\sigma}_m < 0$ if and only if $\sigma_2/\sigma_1 + 2\theta\rho < 0$. In this situation, $\sigma_1 + \bar{\sigma}_m/\theta$ is nonnegative, and we have

$$\frac{1}{\tau}\int_0^\tau \left[\sigma_1 + \frac{\bar{\sigma}_m}{\theta}\left(1 - e^{-\theta x}\right)\right]^2 dx > \left[\sigma_1 + \frac{\bar{\sigma}_m}{\theta}\left(1 - e^{-\theta t}\right)\right]^2.$$

Then it follows

$$\frac{dAV}{d\tau}(\tau) < 0.$$

Therefore, in the case where $\sigma_2/\sigma_1 + 2\theta\rho < 0$, the average variance of the stock return is decreasing in the horizon, implying that the stock is less risky over long horizons. Similarly, we can show the converse case in which with $\sigma_2/\sigma_1 + 2\theta\rho > 0$, the riskiness of the stock increases with the horizon.

Chapter 5

Commodities in Dynamic Asset Allocation: Implications of Mean Reverting Commodity Prices

5.1 Introduction

Commodities have been emerging as an increasingly important class of assets for institutional and individual investors in recent years. Systematic investigation of commodities as an investable asset class goes back at least some 30 years ago [Greer, 1978, Bodie and Rosansky, 1980]. However, the growth of commodity markets to a major alternative investment vehicle is a more recent development. Around 2007, the size of the global commodities derivatives market is estimated to be about 750 billion US dollars [Till and Eagleeye, 2007]. As the markets have grown, more investors have been attracted to commodities. Increased exposure to commodities has been acquired by institutional investors, with pension funds as a notable example, and to a less extent by individual investors as well (see e.g. Mongars and Marchal-Dombrat [2006] and Doyle et al. [2007]).

In the literature of commodity investment, it remains an open question whether and how to include commodities in mainstream portfolios. Studies on commodity investment have generally been based on the performance of investment in commodity futures. The reason is that investment in commodities is mostly by means of derivative products, especially commodity futures, while spot transactions of commodities play little role in commodity investment. Most existing studies on commodity investment apply the one-period mean-variance optimization framework of Markowitz [1952]. In the static mean-variance framework, the key issues investigated by these studies have been whether investment in commodities futures yields a positive risk premium, how such investment covaries with bonds and stocks, and how it hedges against inflation (see e.g. Erb and Harvey [2006], Gordon and Rouwenhorst [2006], Kat and Oomen [2007] and references therein). In this literature, the most controversial issue has presumably been whether or not commodity investment offers a positive risk premium, and if it does, what drives the risk premium. The absence of an appreciably positive risk premium does not necessarily make mean-variance investors refrain from allocating to an asset class, but it will surely make it less attractive or of little practical relevance in most studies. As such, the ongoing debate over the risk premium of commodity futures investment has left it an open question whether or not commodities are an appealing asset class.

Empirical evidence has documented that risk premia in commodity futures markets are timing-varying and predictable. For example, Bessembinder and Chan [1992] show that prices in commodity futures markets can be forecast on the basis of instrumental variables known to possess forecasting power in equity and bond markets. Some studies have found that risk premia of commodity investment vary in different states, like the phase of the business cycle, the stance of monetary policy, market sentiment, and the history of investment returns (see, for example, Jensen et al. [2000, 2002], Wang and Yu [2004], Erb and Harvey [2006], Miffre and Rallis [2007], Nijman and Swinkels [2007], and Vrugt et al. [2007]).

Similarly, time-variation and predictability of asset returns have been well documented in the asset classes of stocks and bonds. What are the implications of time-variation and predictability of asset returns for portfolio choice? For the mainstream asset classes, their implications for portfolio choice have been explored in depth, for example, Kim and Omberg [1996] and Wachter [2002] in the case of stocks. In the case of commodities, however, much less research efforts have been devoted to the implications of time-varying and predictable commodity returns for portfolio decision making. In view of this, this chapter presents a study of the asset class of commodities in an intertemporal framework, with an explicit focus on the time-varying and predictable returns in commodity markets. In the literature of commodity investment, the closest to this study is presumably Hoevenaars et al. [2008], who address the optimal portfolio policy in the context where expected return of alternative assets, including commodity returns, are time-varying and predictable. For technical reasons, however, Hoevenaars et al. [2008] consider only constant-proportion portfolio strategies, and hence abstract from market timing that in principle will arise from return predictability. This study aims to further the understanding of commodity investment, especially in exploiting commodity return predictability by market timing.

This chapter investigates the asset class of commodities in the dynamic optimization framework that was introduced into finance by Merton [1969]. The so-called Merton's problem has been analyzed and extended in various contexts, reflecting different attributes of people's preferences and of financial markets (see, for example, Chapter 9 of Duffie [2001] for a textbook treatment). The literature of dynamic asset allocation, however, has focused predominantly on such traditional asset classes as stocks and bonds. Owing to the growing importance of commodities, it is pertinent to ask, in this established framework, how investors should optimally make their portfolio and consumption decisions when commodities are available in addition to stocks and bonds.

To this end, I introduce into the classical Black-Scholes economy a commodity market. Consistent with the fact that commodity futures are the major commodity investment vehicle, the commodity market is modeled as a futures market. With the addition of the commodity market, the Black-Scholes economy consisting of a riskless bond and a risky stock is augmented to an economy equipped with three asset classes. This three-asset economy, referred to as the "Bond-Stock-Commodity economy" in the following, enables one to capture the richer investment opportunities stemming from the presence of commodities.

The commodity futures market is characterized by a generalized version of the single-factor model in Schwartz [1997]. Following Schwartz [1997], the non-tradable spot commodity price follows a mean-reverting process. Rather than assuming a constant risk premium in the futures market as in Schwartz [1997], I generalize his model by assuming that the risk premium is dependent on the spot commodity price. This generalization can be justified by three reasons. First, as mentioned above, empirical evidence has shown that risk premia in commodity markets are time-varying, and can be predicted by instrumental variables characteristic of the business cycle. Second, empirical studies, for example Fama and French [1988], have identified a strong business cycle component in the variation of spot commodity prices. It suggests that spot commodity prices might have forecast power for risk premia in commodity futures markets. Last but not least, the estimation results of this extended model presented in this chapter has provided strong evidence that the effect of the spot price on the risk premium is significant.

In this simple characterization of the asset class of commodities, the risk premium in the commodity market is predicted by the mean-reverting spot commodity price. As is well known, mean reversion is an important property of commodity prices, and mean reversion of prices has become a prevailing assumption in the literature related to the stochastic behavior of commodity prices, for instance, Gabillon [1995], Schwartz [1997], Geman [2005], to name but a few. Moreover, empirical studies of commodity prices have found evidence of mean reversion to various degrees (for example Bessembinder et al. [1995], Pindyck [2001], and Andersson [2007]).

This study, by relating commodity market returns to spot commodity prices, underscores the implications of the mean-reverting nature of the commodity price for commodity investment and portfolio decisions.

The dynamic framework adopted here makes this study distinct from ones that use a static perspective. In the static one-period paradigm, people are assumed to make a *one-off* investment decision at the beginning of the period in order to maximize their utility over the investment outcome at the end of the period. In comparison, the dynamic framework built on an intertemporal setting allows people to make *intermediate* rebalancing. Undoubtedly the dynamic framework offers a richer structure than the static one does, and arguably it is closer to financial decision-making in practice. It has long been known that unless (i) investors have logarithmic utility, or (ii) the financial market offers a constant investment opportunity set in the sense that both the riskfree rate and the market price of risk are constant, the optimal financial policy derived from the dynamic framework is different from the so-called "myopic" policy based solely on one-period analysis. As will be shown, the time variation and predictability of expected returns in commodity markets, once investigated in the framework of dynamic asset allocation, has profound implications for commodity investment and portfolio decisions, which the static mean-variance analysis is unable to accommodate. Therefore, this study, by virtue of a richer framework, will shed new light on the debate on commodity investment, and enable us to expound on some contentious issues arising from static analysis in light of the findings from a dynamic perspective.

This study contributes to the discussion of commodity investment by taking a novel route to approach the issue. Different from focusing on indices of commodity futures in extant literature, I model commodities into the economy as commodity futures underlying those indices. And this approach may have an advantage in comparison with that based on commodity futures indices. It has been recognized that commodity futures indices embed trading strategies of commodity futures (Gordon and Rouwenhorst [2005], Erb and Harvey [2006]). Owing to differing ways of composition, weighting and rebalancing, different commodity indices imply different trading strategies, and hence may well give divergent pictures of commodity investment returns, even in a common time period. In addition to bringing an element of arbitrariness because of varying ways of index building, the application of indices may blur some important characteristics of commodity investment, like the implications of mean reversion in commodity prices. In contrast, this study directly specifies the underlying commodity futures as such, in the hope of achieving a sharp focus on implications of this property.

I shall consider the optimal financial strategy for an investor in two classical cases. In the first, the investor is concerned with maximizing the expected utility over wealth on some fixed horizon date. The second case I consider is that of an investor who derives utilities over life-time consumption. Of these two cases, the first, terminal wealth case only involves portfolio decisions, and is conceptually easier. The intermediate consumption case, being slightly more complicated, involves both portfolio and consumption decisions. In both cases, the investor is assumed to have constant relative risk aversion, and to be more risk averse than a logarithmic investor.

By the specification of the commodity futures market in this chapter, the risk premium of commodity investment turns out to follow an Ornstein-Uhlenbeck process. This property allows us to approach the dynamic optimization problem in a route similar to that developed by Kim and Omberg [1996] and Wachter [2002], who address the dynamic optimization problem in the context of predictable equity premia. Thanks to the simple structure of the model, the optimal policy and the utility cost of excluding the commodity are solved in closed form. The optimal policy dictates that allocation to commodities is made both for myopic purposes and for intertemporal purposes, whereas stock allocation is made solely out of myopic considerations. The optimal financial strategy involves timing on the spot commodity price, and thus puts forward a theoretical case for the tactical timing strategies studied in some empirical investigations.

The remainder of the chapter is organized as follows. I present the basic model of the Bond-Stock-Commodity economy in the next section. Section 5.3 and 5.4 are devoted to the optimal strategy in the terminal wealth case and in the intermediate consumption case, respectively. In section 5.5, I estimate the model of commodity futures, and offer some representative numerical examples and discussion. Section 5.6 concludes.

5.2 The economy

In the Bond-Stock-Commodity economy, people can invest in three asset classes: bonds, stocks, and commodities. I opt to characterize the traditional asset classes of bonds and stocks by the standard Black and Scholes [1973] model in order to isolate the effect of the introduction of commodities, although it is possible to follow other formulations which were developed in recent years to reflect stochastic interest rates, and the documented predictability of stock returns.

For the riskless bond, the constant interest rate is denoted by r. The stock price S_t follows a geometric Brownian motion

$$dS_t = \mu_1 S_t dt + \sigma_1 S_t dZ_{1,t}, \tag{5.1}$$

where μ_1 and σ_1 are positive constants, and $Z_{1,t}$ is a standard Wiener process.

Following Schwartz [1997], the spot commodity price M_t is specified by the following single-factor model with mean-reverting property:

$$dM_t = \theta(\mu_2 - \ln M_t)M_t dt + \sigma_2 M_t dZ_{M,t}, \tag{5.2}$$

where θ, μ_2, and σ_2 are positive constants, and $Z_{M,t}$ is another standard Wiener process jointly normally distributed with $Z_{1,t}$. Defining $m_t = \ln M_t$, we have

$$dm_t = \theta\left(\mu_2 - \frac{\sigma_2^2}{2\theta} - m_t\right)dt + \sigma_2 dZ_{M,t}. \tag{5.3}$$

As in Schwartz [1997], the spot commodity price is assumed to be non-tradable. Assuming the market price of risk associated with the Wiener process driving the spot commodity price, $Z_{M,t}$, is given by

$$\lambda_{M,t} = \alpha + \beta m_t. \tag{5.4}$$

This assumption is motivated by two empirical findings: (i) expected returns in commodity markets are time-varying and can be predicted by some instrumental variables characteristic of the business cycle; and (ii) there is a strong business cycle component in the variation of spot commodity prices. Note that this model is reduced to the one-factor model of Schwartz [1997] when $\beta = 0$.

Then under the risk-neutral measure

$$dm_t = \left(\widetilde{\mu}_2 - \widetilde{\theta} m_t\right)dt + \sigma_2 d\widetilde{Z}_{M,t},$$

where

$$\widetilde{\theta} := \theta + \sigma_2\beta, \qquad \widetilde{\mu}_2 := \theta\mu_2 - \sigma_2\alpha - \sigma_2^2/2,$$

and $\widetilde{Z}_{M,t}$ is a standard Wiener process under the risk-neutral measure. From the above equation, the distribution of m_T conditional on m_t $(t < T)$ under the risk-neutral measure is normal with mean and variance:

$$\begin{aligned} E_t^{\mathbb{Q}}[m_T] &= e^{-\widetilde{\theta}t}m_t + \left[1 - e^{-\widetilde{\theta}(T-t)}\right]\frac{\widetilde{\mu}_2}{\widetilde{\theta}} \\ \mathrm{Var}_t^{\mathbb{Q}}[m_T] &= \frac{\sigma_2^2}{2\widetilde{\theta}}\left[1 - e^{-2\widetilde{\theta}(T-t)}\right] \end{aligned}$$

From the martingale property of futures prices under the risk-neutral measure, it follows that the futures price of the commodity with maturity T at time t is

$$F_t(T) = E_t^{\mathbb{Q}}[M_T] = \exp\left(E_0^{\mathbb{Q}}[m_T] + \tfrac{1}{2}\mathrm{Var}_0^{\mathbb{Q}}[m_T]\right).$$

Then,

$$F_t(T) = \exp\left\{e^{-\widetilde{\theta}(T-t)}m_t + \left[1 - e^{-\widetilde{\theta}(T-t)}\right]\frac{\widetilde{\mu}_2}{\widetilde{\theta}} + \frac{\sigma_2^2}{4\widetilde{\theta}}\left[1 - e^{-2\widetilde{\theta}(T-t)}\right]\right\}. \qquad (5.5)$$

To facilitate the solution of Merton's problem, I characterize the asset class of commodities by the self-financing portfolio $\mathcal{M}_t$:

$$d\mathcal{M}_t = \mathcal{M}_t\left(r + \sigma_2\lambda_{M,t}\right)dt + \mathcal{M}_t\sigma_2 dZ_{M,t}. \qquad (5.6)$$

This portfolio is formed by a portfolio strategy of the riskless bond and the commodity futures as follows. Hold a long position in the commodity futures, and the futures holding is constantly rolled over to keep the time-to-maturity of futures contracts constant, say equal to ℓ where ℓ is a positive constant. Moreover, the futures holding is such that the notional value of the futures contract (the number of futures contracts times the futures price) at time t is equal to $\mathcal{M}_t e^{\widetilde{\theta}\ell}$. Because futures contracts have zero value, the entire portfolio value $\mathcal{M}_t$ is invested in the riskless bond. For more detail of the portfolio strategy of $\mathcal{M}_t$, see Appendix 5.A.

As such, the specification of the Bond-Stock-Commodity market has been completed. To ease the solution to Merton's problem, however, I reformulate this model by converting the two possibly correlated driving Wiener processes to a *standard* two-dimensional Wiener process (i. e. its two components are independent). Denoting the correlation coefficient between $Z_{1,t}$ and $Z_{M,t}$ by ρ where $|\rho| < 1$, the price dynamics of the two classes of risky assets given by (5.1) and (5.6) can be rewritten as

$$\begin{aligned} dS_t &= \mu_1 S_t dt + \sigma_1 S_t \left[\begin{array}{cc} 1 & 0 \end{array} \right] d\boldsymbol{Z}_t, \\ d\mathcal{M}_t &= \mathcal{M}_t(r + \sigma_2 \lambda_{M,t})dt + \mathcal{M}_t \sigma_2 \left[\begin{array}{cc} \rho & \bar{\rho} \end{array} \right] d\boldsymbol{Z}_t, \end{aligned} \tag{5.7}$$

where $\bar{\rho} := \sqrt{1-\rho^2}$, and $\boldsymbol{Z}_t := \left[\begin{array}{cc} Z_{1,t} & Z_{2,t} \end{array} \right]^\top$ is a standard two-dimensional Wiener process.[1] In terms of the standard vector Wiener process, m_t can be written as

$$dm_t = \theta \left(\mu_2 - \frac{\sigma_2^2}{2\theta} - m_t \right) dt + \sigma_2 \left[\begin{array}{cc} \rho & \bar{\rho} \end{array} \right] d\boldsymbol{Z}_t.$$

Given the specification of the financial market as in (5.7) and the constant interest rate r, the market price of risk associated with the standard vector Wiener process $\boldsymbol{Z}_t$ is

$$\boldsymbol{\lambda}_t = \left[\begin{array}{c} \lambda_1 \\ \lambda_{2,t} \end{array} \right],$$

where

$$\begin{aligned} \lambda_1 &= \frac{\mu_1 - r}{\sigma_1}, \\ \lambda_{2,t} &= \lambda_2(m_t) := \frac{\lambda_{M,t}}{\bar{\rho}} - \frac{\rho}{\bar{\rho}}\lambda_1. \end{aligned} \tag{5.8}$$

That is to say, the market price of risk with respect to $Z_{1,t}$ is constant, whereas that with respect to $Z_{2,t}$ is stochastic and dependent on the commodity price. As such, the investment opportunity set is stochastic in the Bond-Stock-Commodity economy, and the optimal strategy should be different from the myopic one unless the utility function of the investor is logarithmic. Moreover, being a linear transform of m_t, $\lambda_{2,t}$ also follows an Ornstein-Uhlenbeck process:

$$d\lambda_{2,t} = \theta \left(\bar{\lambda}_2 - \lambda_{2,t} \right) dt + \frac{\beta\sigma_2}{\bar{\rho}} \left[\begin{array}{cc} \rho & \bar{\rho} \end{array} \right] d\boldsymbol{Z}_t, \tag{5.9}$$

where

$$\bar{\lambda}_2 = \frac{\beta}{\bar{\rho}} \left(\mu_2 - \frac{\sigma_2^2}{2\theta} \right) + \frac{\alpha}{\bar{\rho}} - \frac{\rho}{\bar{\rho}}\lambda_1.$$

For a market to exclude arbitrage, it suffices that the Novikov condition holds (see, for example, Chapter 6 in Duffie [2001]):

$$E\left[\exp\left(\frac{1}{2} \int_0^T \boldsymbol{\lambda}_t^\top \boldsymbol{\lambda}_t dt \right) \right] < \infty.$$

---

[1] In this chapter, boldface notation is used to denote vectors and matrices.

It can be verified that the Novikov condition holds indeed in our model,[2] so the Bond-Stock-Commodity economy is free of arbitrage. Moreover, this economy can be shown to be a complete market (see e.g. Kreps and Pliska [1981]), with a unique state-price density ξ_t given by

$$\frac{d\xi_t}{\xi_t} = -rdt - \boldsymbol{\lambda}_t^\top d\boldsymbol{Z}_t, \quad \text{and} \quad \xi_0 = 1. \tag{5.10}$$

Now turn to an investor with initial wealth W_0. Her consumption plan is characterized by an consumption-rate process c_t, and her portfolio plan is a process of portfolio weights in the two risky assets $\boldsymbol{x}_t = \begin{bmatrix} x_{S,t} & x_{M,t} \end{bmatrix}^\top$, where $x_{S,t}$ and $x_{M,t}$ denote the portfolio weights in the stock, S_t, and in the investable representative commodity, $\mathcal{M}_t$, respectively. The residual, $1 - x_{S,t} - x_{M,t}$, is allocated to the riskless bond.

From the self-financing property of the consumption-portfolio plan, it follows that the wealth process W_t is given by

$$dW_t = W_t \left[r + \boldsymbol{x}_t^\top \boldsymbol{\sigma} \boldsymbol{\lambda}_t\right] dt - c_t dt + W_t \boldsymbol{x}_t^\top \boldsymbol{\sigma} d\boldsymbol{Z}_t, \tag{5.11}$$

where

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_1 & 0 \\ \sigma_2 \rho & \sigma_2 \bar{\rho} \end{bmatrix}.$$

Note that given $x_{M,t}$, we can calculate the corresponding holding of the underlying commodity futures contracts as follows. Recall that in one unit of the representative commodity $\mathcal{M}_t$, the holding of the futures contract with constant time to maturity ℓ has a notional value of $\mathcal{M}_t e^{\widetilde{\theta}\ell}$. When $x_{M,t}$ weight of total wealth is allocated to $\mathcal{M}_t$, it implies the ratio of the notional value of the underlying future contract to the wealth value W_t is $x_{M,t} e^{\widetilde{\theta}\ell}$.

In the terminal wealth case, the investor solves the following dynamic optimization problem:

$$\begin{aligned} &\sup_{\boldsymbol{x}_t} E\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \\ \text{s.t.} \quad dW_t = &W_t \left[r + \boldsymbol{x}_t^\top \boldsymbol{\sigma} \boldsymbol{\lambda}_t\right] dt + W_t \boldsymbol{x}_t^\top \boldsymbol{\sigma} d\boldsymbol{Z}_t, \end{aligned} \tag{5.12}$$

where γ is the constant rate of relative risk aversion. For reasons that will become clear later, γ is assumed to be larger than one throughout the chapter to ensure the existence of a well-behaved solution. This assumption is empirically relevant as it is generally

[2] As λ_1 is constant and hence satisfies the Novikov condition, we only have to verify that it is also true for the Ornstein-Uhlenbeck $\lambda_{2,t}$. Dokuchaev [2007] has proved that a market price of risk following an Ornstein-Uhlenbeck process satisfies the Novikov condition. Following the same reasoning as in Dokuchaev [2007], we can prove that this condition applies in our model.

supported by empirical studies of people's risk aversion (see e.g. Friend and Blume [1975], Pindyck [1988], and Szpiro [1986]), and by the literature on the equity premium puzzle. In the intermediate consumption case, the dynamic optimization problem is

$$\begin{aligned} &\sup_{\boldsymbol{x}_t, c_t} E\left[\int_0^T e^{-\eta t}\frac{c_t^{1-\gamma}}{1-\gamma}dt\right] \\ \text{s.t.} \quad dW_t =& W_t\left[r+\boldsymbol{x}_t^\top \boldsymbol{\sigma}\boldsymbol{\lambda}_t\right]dt - c_t dt + W_t \boldsymbol{x}_t^\top \boldsymbol{\sigma} d\boldsymbol{Z}_t, \\ &W_T \geq 0, \end{aligned} \tag{5.13}$$

where η denotes the subjective discount rate.

5.3 Pure portfolio optimization

I start with the terminal wealth case, in which there are only portfolio decisions to make. Using the martingale method, the dynamic optimization problem (5.12) is equivalent to the following static variational problem [Cox and Huang, 1991]:

$$\begin{aligned} &\sup_{\boldsymbol{x}_t} E\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \\ \text{s.t.} \quad &W_0 = E\left[\xi_T W_T\right] \end{aligned} \tag{5.14}$$

In particular, the budget constraint in (5.12) is equivalent to the static one in (5.14) that is formulated in terms of the unique state price density. For a solution to this static optimization problem to exist, it suffices that $E\left[\xi_T^{-1}\right]$ is finite [Cox and Huang, 1991], namely the growth-optimal portfolio has a finite expectation. Under the assumption that this condition applies, the optimal terminal wealth is determined by

$$W_T^* = (k\xi_T)^{-1/\gamma},$$

where k is a Lagrange multiplier determined by substituting the optimal terminal wealth into the static budget constraint.

Following Cox and Huang [1989], I define a new variable

$$N_t = (k\xi_t)^{-1}.$$

By Ito's formula,

$$\frac{dN_t}{N_t} = \left(r+\lambda_1^2+\lambda_{2,t}^2\right)dt + \left[\begin{array}{cc} \lambda_1 & \lambda_{2,t} \end{array}\right] d\boldsymbol{Z}_t. \tag{5.15}$$

From the definition of N_t and the fact that $\xi_t W_t$ is a martingale, we have

$$W_t = \frac{1}{\xi_t} E_t\left[\xi_T W_T^*\right] = \frac{1}{k\xi_t} E_t\left[k\xi_T (k\xi_T)^{-\frac{1}{\gamma}}\right] = N_t E\left[N_T^{\frac{1}{\gamma}-1} \Big| \lambda_{2,t}, N_t\right],$$

where the last equality follows from the fact that $\lambda_{2,t}$ and N_t together form a strong Markov process, and hence $\lambda_{2,t}$ and N_t are all the investor needs to know to evaluate moments of N_T at time t. Therefore, we can define

$$W_t := F(N_t, \lambda_{2,t}, t; T).$$

5.3.1 Optimal wealth

To simplify notation, I define the following parameters

$$\begin{aligned} a_1 &= 2\left(\frac{1-\gamma}{\gamma}\beta\sigma_2 - \theta\right), \quad a_2 = \frac{1}{\gamma}\left(\frac{\beta\sigma_2}{\bar{\rho}}\right)^2, \\ q &= \sqrt{a_1^2 - 4\frac{1-\gamma}{\gamma}a_2}, \qquad \lambda^* = \theta\bar{\lambda}_2 + \frac{1-\gamma}{\gamma}\frac{\rho}{\bar{\rho}}\beta\sigma_2\lambda_1. \end{aligned} \tag{5.16}$$

Then the optimal wealth can be presented as follows.

Lemma 5.3.1. *For an investor concerned with maximizing the expected utility over wealth at time T as described in (5.12), the optimal wealth is given by*

$$W_t = F(N_t, \lambda_{2,t}, t; T) = N_t^{\frac{1}{\gamma}} H(\lambda_{2,t}, T-t), \tag{5.17}$$

where

$$H(\lambda_{2,t}, \tau) = \exp\left\{\frac{1}{\gamma}\left[\frac{1}{2}A_1(\tau)\lambda_{2,t}^2 + A_2(\tau)\lambda_{2,t} + A_3(\tau)\right]\right\}, \tag{5.18}$$

and

$$\begin{aligned} A_1(\tau) &= \frac{1-\gamma}{\gamma}\frac{2\left(1-e^{-q\tau}\right)}{2q - (q+a_1)\left(1-e^{-q\tau}\right)}, \\ A_2(\tau) &= \frac{1-\gamma}{\gamma}\frac{4\lambda^*\left(1-e^{-q\tau/2}\right)^2}{q\left[2q - (q+a_1)\left(1-e^{-q\tau}\right)\right]}, \\ A_3(\tau) &= \int_0^\tau \left[\frac{a_2}{2}A_2^2(x) + \lambda^* A_2(x) + \frac{\gamma a_2}{2}A_1(x) + (1-\gamma)r + \frac{1-\gamma}{2\gamma}\lambda_1^2\right] dx. \end{aligned} \tag{5.19}$$

Proof. Given that $W_t = F(N_t, \lambda_{2,t}, t; T)$ and the stochastic differential equations (5.9) and (5.15) for $\lambda_{2,t}$ and N_t, we can write the wealth process in the form of a stochastic differential equation by applying Ito's formula:

$$dW_t = \mu_W dt + \sigma_W d\boldsymbol{Z}_t, \tag{5.20}$$

where

$$\begin{aligned}\mu_W &= \frac{\partial F}{\partial t} + \frac{\partial F}{\partial N} N_t \left(r + \lambda_1^2 + \lambda_{2,t}^2\right) + \frac{\partial F}{\partial \lambda_2}\theta\left(\bar{\lambda}_2 - \lambda_{2,t}\right) + \frac{1}{2}\frac{\partial^2 F}{\partial \lambda_2^2}\left(\frac{\beta\sigma_2}{\bar{\rho}}\right)^2 \\ &\quad + \frac{1}{2}\frac{\partial^2 F}{\partial N^2} N_t^2(\lambda_1^2 + \lambda_{2,t}^2) - \frac{\partial^2 F}{\partial \lambda_2 \partial N} N_t \frac{\beta\sigma_2}{\bar{\rho}}(\rho\lambda_1 + \bar{\rho}\lambda_{2,t}), \\ \sigma_W &= \frac{\partial F}{\partial N} N_t [\ \lambda_1 \quad \lambda_{2,t}\ ] + \frac{\partial F}{\partial \lambda_2}\frac{\beta\sigma_2}{\bar{\rho}}[\ \rho \quad \bar{\rho}\ ].\end{aligned}$$

Because W_t is a self-financing wealth process, no arbitrage requires

$$\mu_W - rF = \sigma_W \boldsymbol{\lambda}_t.$$

Writing it explicitly leads to the following partial differential equation (PDE)

$$\begin{aligned}&\frac{\partial F}{\partial t} + r\frac{\partial F}{\partial N} N_t + \left(\theta\bar{\lambda}_2 - \theta\lambda_{2,t} - \frac{\beta\sigma_2}{\bar{\rho}}\rho\lambda_1 - \beta\sigma_2\lambda_{2,t}\right)\frac{\partial F}{\partial \lambda_2} + \frac{1}{2}\left(\frac{\beta\sigma_2}{\bar{\rho}}\right)^2 \frac{\partial^2 F}{\partial \lambda_2^2} \\ &\qquad + \frac{1}{2}\frac{\partial^2 F}{\partial N^2} N_t^2(\lambda_1^2 + \lambda_{2,t}^2) + \frac{\beta\sigma_2}{\bar{\rho}}\frac{\partial^2 F}{\partial \lambda_2 \partial N} N_t(\rho\lambda_1 + \bar{\rho}\lambda_{2,t}) = rF. \quad (5.21)\end{aligned}$$

F also satisfies the boundary condition,

$$F(N_T, \lambda_{2,T}, T; T) = W_T^*.$$

It is noteworthy that this PDE bears a close resemblance to the PDE for the optimal wealth process in [Wachter, 2002, Eq. (20)], where the optimal portfolio choice problem is addressed in the context of mean-reverting stock risk premia. The resemblance arises from the fact that the market price of risk in Wachter's model is characterized by an Ornstein-Uhlenbeck process, as is $\lambda_{2,t}$ in the Bond-Stock-Commodity model.

The PDE can be solved by first guessing a general form for the solution. Enlightened by the solution to the PDE in Wachter [2002], I guess the form given by (5.17) and (5.18). Substituting them back into (5.21) yields a quadratic equation for $\lambda_{2,t}$; from the fact that both the constant term and the coefficients on $\lambda_{2,t}^2$ and $\lambda_{2,t}$ must be zero, one obtains a system of three ordinary differential equations:

$$\begin{aligned}\frac{dA_1}{d\tau}(\tau) &= a_2 A_1^2(\tau) + a_1 A_1(\tau) + \frac{1-\gamma}{\gamma}, \\ \frac{dA_2}{d\tau}(\tau) &= a_2 A_1(\tau)A_2(\tau) + \left(\frac{a_1}{2}\right)A_2(\tau) + \lambda^* A_1(\tau), \qquad (5.22) \\ \frac{dA_3}{d\tau}(\tau) &= \frac{a_2}{2}A_2^2(\tau) + \lambda^* A_2(\tau) + \frac{\gamma a_2}{2}A_1(\tau) + (1-\gamma)r + \frac{1-\gamma}{2\gamma}\lambda_1^2.\end{aligned}$$

Equations of the same form appear in Kim and Omberg [1996] and Wachter [2002], and the solution method is standard. Following Wachter [2002], I assume that $\gamma > 1$ to ensure the existence of a well-behaved solution. Under this assumption, the solution is given by (5.19).

For the validity of the optimal solution (5.17), some technical conditions need to be satisfied [Cox and Huang, 1989]. Appendix 5.A verifies that these conditions hold. Therefore the optimal wealth is given by (5.17). □ □

5.3.2 Optimal portfolio plan

Turn to the optimal portfolio plan, a plan that secures the optimal wealth. In the martingale solution, the optimal portfolio plan can be obtained by equating the diffusion terms in the two characterizations of the optimal wealth given by (5.11) and (5.20). So the optimal strategy can be summarized as follows.

Theorem 5.3.2. *For an investor facing the problem (5.12) in the Bond-Stock-Commodity economy, the optimal portfolio plan* $\boldsymbol{x}_t^* := \begin{bmatrix} x_{S,t}^* & x_{M,t}^* \end{bmatrix}^\top$ *is given by*

$$\begin{aligned}\boldsymbol{x}_t^* &= \frac{1}{W_t}\left(\boldsymbol{\sigma}^\top\right)^{-1}\sigma_W^\top \\ &= \underbrace{\frac{1}{\gamma}\begin{bmatrix} \frac{\lambda_1}{\sigma_1} - \frac{\rho}{\sigma_1\bar{\rho}}\lambda_{2,t} \\ \frac{1}{\sigma_2\bar{\rho}}\lambda_{2,t} \end{bmatrix}}_{\textit{myopic part}} + \underbrace{\frac{1}{\gamma}\begin{bmatrix} 0 \\ \frac{\beta}{\bar{\rho}}\left[A_1(T-t)\lambda_{2,t} + A_2(T-t)\right] \end{bmatrix}}_{\textit{intertemporal part}}. \end{aligned} \tag{5.23}$$

As is standard in the literature, the optimal strategy in the above presentation is decomposed into two parts: a myopic part, and an intertemporal part. The myopic part, independent of investment horizon, is the allocation that an investor would choose if she ignored changes in the investment opportunity set or if her utility function is logarithmic. The interpretation from the perspective of logarithmic utility can be seen directly by setting γ to one: when γ is one, A_1 and A_2 are zero, and the second part disappears. The intertemporal allocation, the concept of which was first introduced by Merton [1971] and repeated in many subsequent studies, depends on the investment horizon and stems from stochastic variations in the investment opportunity set.

Let us look at the allocation to the two risky assets in more detail. Given the empirical evidence presented in Section 5.5 that the risk premium in the commodity market is decreasing in the spot commodity price, namely a significantly negative estimate of β in (5.4), it is assumed that $\beta < 0$ in the following discussion. The *stock allocation*,

$x^*_{S,t}$, consists solely of a myopic part. It should not come as a surprise, considering that the stochastic changes in the investment opportunity set in the Bond-Stock-Commodity economy are caused by the variations of the commodity price as shown in (5.8), and hence the commodity should be in a better position to deal with them. The stock allocation, written as a combination of two terms $\frac{\lambda_1}{\gamma\sigma_1} - \frac{\rho}{\gamma\sigma_1\bar{\rho}}\lambda_{2,t}$, has a natural economic interpretation. The first term, $\frac{\lambda_1}{\gamma\sigma_1}$, is the classical stock allocation in the Black-Scholes economy [Merton, 1969]. The second term, $-\frac{\rho}{\gamma\sigma_1\bar{\rho}}\lambda_{2,t}$, is more interesting for our purposes, as it arises from the introduction of the commodity. Because of the relationship between $\lambda_{2,t}$ and the commodity price as given in (5.4) and (5.8), the second term implies that the stock allocation is dependent on the commodity price. And the dependence may take three forms, according to the way the commodity and the stock covary with each other: (i) when the stock price is positively correlated with the commodity price ($\rho > 0$), the stock weight is increasing with the commodity price; (ii) when the stock price is negatively correlated with the commodity price ($\rho < 0$), the stock weight is decreasing with the commodity price; and (iii) when they are independent from each other ($\rho = 0$), the stock weight is immune to the commodity price variation, and constant at $\frac{\lambda_1}{\gamma\sigma_1}$.

Different from the case for the stock, the *commodity allocation* $x^*_{M,t}$ is made both for myopic purposes and for intertemporal purposes. First consider the *myopic demand for the commodity*, $\lambda_{2,t}/\gamma\sigma_2\bar{\rho}$. With $\beta < 0$, it is a decreasing function of the commodity price, and the myopic demand requires to sell the commodity when its price rises, and to buy when its price drops. This property follows from the fact that the instantaneous expected return on the commodity is negatively related to the current commodity price. Whether the investor should be long or short the commodity depends on the sign of $\lambda_{2,t}$. Setting $\lambda_{2,t} = 0$, we can solve the threshold value of the commodity price for the myopic demand

$$\bar{M}^{mpc} = \exp\left(\frac{\rho\lambda_1 - \alpha}{\beta}\right).$$

With $\beta < 0$, we can distinguish three cases: (i) when the commodity price is lower than $\bar{M}^{mpc}$, and then $\lambda_{2,t}$ is positive, the myopic demand is a long position; (ii) when the commodity price is greater than $\bar{M}^{mpc}$, and then $\lambda_{2,t}$ is negative, the investor is short the commodity; and (iii) when the commodity price is equal to $\bar{M}^{mpc}$, and then $\lambda_{2,t}$ is zero, the optimal policy dictates no exposure to the commodity for myopic purposes. These are natural results if recalling that the myopic allocation is concerned only with instantaneous returns of assets.

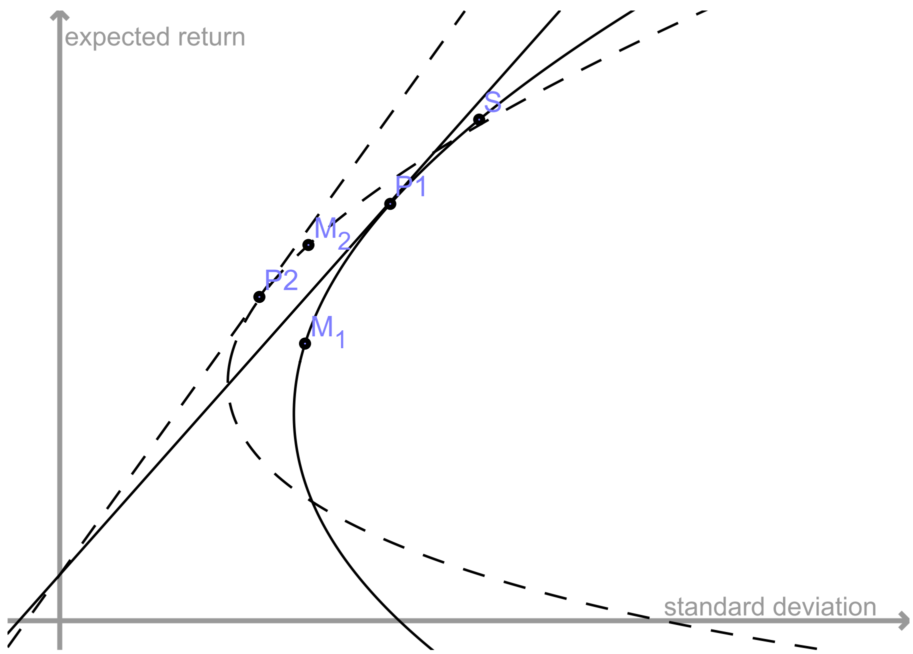

Figure 5.1: **The dependence of the myopic allocation on the commodity price: an illustration** The figure shows the instantaneous mean-variance optimization for two different commodity price M_1 and M_2. The corresponding tangency portfolios are "P1" and "P2".

Thus, the myopic demand both for the stock and for the commodity is dependent on the commodity price. This dependence can be accounted for more intuitively by analogy with mean-variance optimization. As is well known (see, for example, Chapter 13 in Ingersoll [1987]), optimal myopic allocation to risky assets, i.e. the portfolio of risky assets optimally chosen by log investors, can be interpreted as the tangency portfolio in the *instantaneous* standard deviation-expectation graph as illustrated in Figure 5.1. In the Bond-Stock-Commodity economy, the expectation and variance of the return on the stock is fixed, so the stock is characterized by a fixed point in the figure. However, the expected return on the commodity is conditional on the current commodity price, so the locus of the commodity in the standard deviation-expectation graph is time-varying. Suppose that the spot commodity price is M_1 at a certain time, a myopic investor would find the optimal allocation by looking for the tangency portfolio (labeled as "P1" in the figure), a portfolio based on the fixed locus of the stock and the current locus of the commodity. If the spot commodity price changes to M_2, say, and the commodity changes to a corresponding new locus, then the corresponding new tangency portfolio (labeled as "P2") will be formed according to the "new" commodity. As such, the myopic allocation changes with the variation of the commodity price.

Now turn to the *intertemporal demand for the commodity*. From its expression

$$\frac{\beta}{\gamma\bar{\rho}}\left[A_1(T-t)\lambda_{2,t} + A_2(T-t)\right], \tag{5.24}$$

it follows that the intertemporal demand changes with the spot commodity price, through the presence of $\lambda_{2,t}$, and with the investment horizon, through the presence of $A_1(T-t)$ and $A_2(T-t)$. First look at the impact of the spot commodity price. Because $\beta < 0$ and $A_1(\tau)$ is negative (see Appendix 5.A), the intertemporal allocation is a decreasing function of the commodity price. In other words, as in the myopic allocation to commodity, the intertemporal allocation is to buy the commodity when its price falls, and to sell when its price rises. By definition, the intertemporal allocation is concerned with returns on assets *beyond* the next period (an infinitesimal period in continuous time). For the mean-reverting commodity, a price hike implies that not only the return on the commodity over the next infinitesimal period is getting worse[3], but the returns beyond the next period are deteriorating as well. And it calls for, as a response, lowering the exposure to the commodity beyond what is done is the myopic allocation. Conversely, a price slump implies improved prospects of future returns, and requires an increased exposure in the intertemporal allocation. These observations may help to understand why the intertemporal demand is decreasing in the commodity price.

Consider now the dependence of the intertemporal allocation on the investment horizon. In particular, should a long-term investor allocate more to the commodity than a short-term one? The horizon effect of the intertemporal allocation to the commodity also represents the entire horizon effect of the total risky allocation, as it is the sole element dependent on the horizon. The horizon effect can be characterized by

$$\frac{\beta}{\gamma\bar{\rho}}\left(A_1'(T-t)\lambda_{2,t} + A_2'(T-t)\right), \tag{5.25}$$

which follows from differentiating (5.24) with respect to the horizon. As shown in Appendix 5.A, $A_1'(\tau)$ is negative, whereas $A_2'(\tau)$ can be positive or negative, depending on the sign of λ^*, where λ^*, as defined in (5.16), includes the parameters characterizing the financial market and the risk aversion of the investor. Thus, without imposing further constraints on the parameter values, we cannot decide the sign of (5.25), or the horizon effect. Further discussion of horizon effect will be given in the numerical example in Section 5.5.

[3]This implication has been captured by the myopic allocation.

After looking into its two components, we are ready to consider the *total commodity allocation*. First of all, the total commodity allocation is decreasing in the commodity price since both components are decreasing in the price. Another question that one may ask is when the investor should be long or short the commodity as a whole. To answer this question, one can derive the following threshold commodity price for deciding the long/short position of the total commodity allocation

$$\bar{M}^{ttl} = \exp\left[\frac{\rho\lambda_1 - \alpha}{\beta} - \frac{\sigma_2\bar{\rho}A_2(T-t)}{1+\sigma_2\beta A_1(T-t)}\right]. \tag{5.26}$$

Therefore, when the commodity price is lower, or higher than this threshold price, the investor should be long, or short the commodity, respectively. The threshold $\bar{M}^{ttl}$ depends on the investment horizon, so it is possible that other things being equal, one investor is short the commodity and another is long simply because they have different investment horizons. Unless $\lambda^* = 0$, the threshold in the total commodity allocation, $\bar{M}^{ttl}$, is different from that in the myopic commodity allocation, $\bar{M}^{mpc}$. This highlights the difference between the dynamic framework and the static one: while a myopic investor would be short the commodity, the optimal allocation may be a long position if intertemporal rebalancing is allowed. This difference is attributed to the intertemporal allocation to the commodity, in which the following threshold of deciding a long/short position is used:

$$\bar{M}^{int} = \exp\left[\frac{\rho\lambda_1 - \alpha}{\beta} - \frac{\bar{\rho}A_2(T-t)}{\beta A_1(T-t)}\right]. \tag{5.27}$$

This threshold is different from $\bar{M}^{mpc}$ unless $\lambda^* = 0$, and it may be less or greater than $\bar{M}^{mpc}$, depending on the sign of λ^*.

5.3.3 The importance of commodities as an asset class: welfare analysis

For an emerging asset class like commodities, it is natural to ask how important it is to take it into account when making investment decisions. In other words, how costly is it if the new asset class is left out in investment decision-making? To address this question, welfare analysis is applied, as is standard in the literature. In this chapter, the importance of incorporating the new asset class, or the utility cost of omitting it, is measured by the percentage extra initial wealth that is necessary to bring the investor to the same expected utility as is obtained by following the optimal strategy. The utility cost can be solved in closed form.

Proposition 5.3.3. *Suppose that in the Bond-Stock-Commodity economy, an investor is concerned with maximizing the expected utility over wealth at time T as described in (5.12). If the commodity is excluded in portfolio decisions, then L percent of extra initial wealth is needed to achieve the same expected utility level as is obtained by following the optimal strategy (5.23), and*

$$\frac{L}{100} = \exp\left\{\frac{1}{1-\gamma}\left[\frac{1}{2}A_1(T)\lambda_{2,0}^2 + A_2(T)\lambda_{2,0} + \int_0^T\left(\frac{a_2}{2}A_2^2(x) + \lambda^* A_2(x) + \frac{\gamma a_2}{2}A_1(x)\right)dx\right]\right\} - 1. \quad (5.28)$$

Proof. First, we need to know the expected utility from the optimal strategy, namely the indirect utility function. Cox and Huang [1989] show that the indirect utility function $J(W_t, \lambda_{2,t}, t)$ satisfies the differential equation

$$\frac{\partial J}{\partial W} = \frac{1}{N_t}.$$

From (5.17), it follows that

$$\frac{1}{N_t} = W_t^{-\gamma} H(\lambda_{2,t}, T-t)^\gamma.$$

Then the differential equation becomes

$$\frac{\partial J}{\partial W} = W_t^{-\gamma} H(\lambda_{2,t}, T-t)^\gamma.$$

Therefore, the boundary condition $J(W_T, \lambda_{2,T}, T) = W_T^{1-\gamma}/(1-\gamma)$ implies that the indirect utility function is

$$J(W_t, \lambda_{2,t}, t) = \frac{W_t^{1-\gamma}}{1-\gamma} H(\lambda_{2,t}, T-t)^\gamma.$$

When the commodity is left out, the Bond-Stock-Commodity economy is reduced to the standard Black-Scholes economy. It is well known that in this specification, the indirect utility is

$$J^{BS}(W_t, t) = \frac{W_t^{1-\gamma}}{1-\gamma}\exp\left[\left(r(1-\gamma) + \frac{1}{2}\frac{1-\gamma}{\gamma}\lambda_1^2\right)(T-t)\right].$$

From the definition of L

$$J^{BS}\left(\frac{100+L}{100}W_0, 0\right) = J(W_0, \lambda_{2,0}, 0),$$

(5.28) follows immediately. $\square$ $\square$

The exponent in the right hand side of (5.28) is a quadratic function of the initial commodity price. From the quadratic form, it follows that with either very high or very low commodity prices, the welfare loss is relatively large, whereas with intermediate commodity prices, the loss is relatively small. The effect on the welfare loss of other factors, like the investment horizon and the risk aversion of the investor, will be discussed in the numerical illustrations in Section 5.5.

5.4 Optimal portfolio and consumption decisions

Now consider the case where the investor derives utility from intermediate consumption. In this case, apart from deciding what asset mix to hold, the investor needs to decide what fraction of wealth to consume. Thus, assuming utility over consumption allows Merton's problem to be related to people's financial decisions in a way that the previous terminal wealth case does not. On the other hand, the intermediate-consumption case has a close link with the terminal wealth case, in that the the single optimization problem in the former case can be thought of as a series of optimization problems for a continuum of future dates [Wachter, 2002]. In particular, the investor with utility over consumption decides the optimal series of consumption events, and then applies the terminal wealth analysis to each future consumption event. This is analogous to the equivalence between a bond that pays coupon continuously and a continuum of zero-coupon bonds. In the following martingale solution, I shall use this insight and follow a procedure similar to Wachter [2002].

The martingale approach transforms the budget constraint in (5.13) into a static one,

$$W_0 = E\left[\int_0^T \xi_t c_t dt\right]. \tag{5.29}$$

The optimal consumption plan follows from the first order condition of the optimization problem in static form,[4]

$$c_t^\star = (\mathcal{K}\xi_t)^{-\frac{1}{\gamma}} e^{-\frac{1}{\gamma}\eta t}, \tag{5.30}$$

where $\mathcal{K}$ is a Lagrange multiplier determined by inserting $c_t^\star$ into (5.29).

The optimal portfolio plan is determined so as to meet the need to finance the consumption plan (5.30). The wealth at time t, denoted by $\mathcal{W}_t$, is the discounted value of

[4] Here we work under the same technical assumption that $E(\xi_T^{-1})$ is finite as in the terminal wealth case. The other technical conditions for the solution's validity are proved in Appendix 5.A.

future consumption till time T,

$$\mathcal{W}_t = \frac{1}{\xi_t} E_t \left[\int_t^T \xi_s c_s^\star ds \right].$$

As in the terminal wealth case, define a new variable

$$\mathcal{N}_t = (\mathcal{K}\xi_t)^{-1}.$$

It follows from Ito's formula that

$$\frac{d\mathcal{N}_t}{\mathcal{N}_t} = \left(r + \lambda_1^2 + \lambda_{2,t}^2\right) dt + \left[\begin{array}{cc} \lambda_1 & \lambda_{2,t} \end{array} \right] d\boldsymbol{Z}_t. \tag{5.31}$$

That is, $\mathcal{N}_t$ has the same dynamics as N_t, but with a different initial value. From the introduction of $\mathcal{N}_t$, and the strong Markov property of $\left[\begin{array}{cc} \mathcal{N}_t & \lambda_{2,t} \end{array} \right]^\top$, it follows that

$$\mathcal{W}_t = \mathcal{N}_t E \left[\int_t^T \mathcal{N}_s^{\frac{1}{\gamma}-1} e^{-\frac{1}{\gamma}\eta s} ds \middle| \lambda_{2,t}, \mathcal{N}_t \right].$$

Therefore one can define

$$\mathcal{W}_t := \mathcal{G}(\mathcal{N}_t, \lambda_{2,t}, t; T).$$

5.4.1 Optimal wealth

The optimal wealth assuming interim consumption can be characterized as follows.

Lemma 5.4.1. *For an investor concerned with maximizing the expected utility over life-time consumption as described in (5.13), the optimal wealth is given by*

$$\mathcal{W}_t = \mathcal{G}(\mathcal{N}_t, \lambda_{2,t}, t; T) = \mathcal{N}_t^{\frac{1}{\gamma}} e^{-\frac{\eta}{\gamma}t} \int_t^T \mathcal{H}(\lambda_{2,t}, s-t) ds, \tag{5.32}$$

where

$$\mathcal{H}(\lambda_{2,t}, s-t) = \exp\left\{ \frac{1}{\gamma} \left[\tfrac{1}{2}\mathcal{A}_1(s-t)\lambda_{2,t}^2 + \mathcal{A}_2(s-t)\lambda_{2,t} + \mathcal{A}_3(s-t) \right] \right\}, \tag{5.33}$$

and

$$\begin{aligned} \mathcal{A}_1(\tau) =& A_1(\tau), \\ \mathcal{A}_2(\tau) =& A_2(\tau), \\ \mathcal{A}_3(\tau) =& A_3(\tau) - \eta\tau \end{aligned} \tag{5.34}$$

Proof. Applying Ito's formula to $\mathcal{W}_t = \mathcal{G}(\mathcal{N}_t, \lambda_{2,t}, t; T)$, one has

$$d\mathcal{W}_t = \mu_{\mathcal{W}} dt + \sigma_{\mathcal{W}} d\boldsymbol{Z}_t, \tag{5.35}$$

where

$$\begin{aligned} \mu_{\mathcal{W}} &= \frac{\partial \mathcal{G}}{\partial t} + \frac{\partial \mathcal{G}}{\partial \mathcal{N}} \mathcal{N}_t \left(r + \lambda_1^2 + \lambda_{2,t}^2\right) + \frac{\partial \mathcal{G}}{\partial \lambda_2} \theta \left(\bar{\lambda}_2 - \lambda_{2,t}\right) + \frac{1}{2} \frac{\partial^2 \mathcal{G}}{\partial \lambda_2^2} \left(\frac{\beta \sigma_2}{\bar{\rho}}\right)^2 \\ &\quad + \frac{1}{2} \frac{\partial^2 \mathcal{G}}{\partial \mathcal{N}^2} \mathcal{N}_t^2 (\lambda_1^2 + \lambda_{2,t}^2) - \frac{\partial^2 \mathcal{G}}{\partial \lambda_2 \partial \mathcal{N}} \mathcal{N}_t \frac{\beta \sigma_2}{\bar{\rho}} (\rho \lambda_1 + \bar{\rho} \lambda_{2,t}), \\ \sigma_{\mathcal{W}} &= \frac{\partial \mathcal{G}}{\partial \mathcal{N}} \mathcal{N}_t [\ \lambda_1 \quad \lambda_{2,t}\ ] + \frac{\partial \mathcal{G}}{\partial \lambda_2} \frac{\beta \sigma_2}{\bar{\rho}} [\ \rho \quad \bar{\rho}\ ]. \end{aligned}$$

Different from the terminal wealth case, the portfolio process $\mathcal{W}_t$ is not self-financing, since a continuous consumption flow $c_t^\star$ is withdrawn. Thus $\mathcal{G}(\mathcal{N}_t, \lambda_{2,t}, t)$ itself does not satisfy the generalized Black-Scholes equation. Instead, in this case no arbitrage requires

$$\mu_{\mathcal{W}} + c_t^\star - r\mathcal{G} = \sigma_{\mathcal{W}} \boldsymbol{\lambda}_t.$$

Writing it explicitly, we have the following PDE for $\mathcal{G}$,

$$\begin{aligned} &\frac{\partial \mathcal{G}}{\partial t} + r \frac{\partial \mathcal{G}}{\partial \mathcal{N}} \mathcal{N}_t + \left(\theta \bar{\lambda}_2 - \theta \lambda_{2,t} - \frac{\beta \sigma_2}{\bar{\rho}} \rho \lambda_1 - \beta \sigma_2 \lambda_{2,t}\right) \frac{\partial \mathcal{G}}{\partial \lambda_2} + \frac{1}{2} \left(\frac{\beta \sigma_2}{\bar{\rho}}\right)^2 \frac{\partial^2 \mathcal{G}}{\partial \lambda_2^2} \\ &\quad + \frac{1}{2} \frac{\partial^2 \mathcal{G}}{\partial \mathcal{N}^2} \mathcal{N}_t^2 (\lambda_1^2 + \lambda_{2,t}^2) + \frac{\beta \sigma_2}{\bar{\rho}} \frac{\partial^2 \mathcal{G}}{\partial \lambda_2 \partial \mathcal{N}} \mathcal{N}_t (\rho \lambda_1 + \bar{\rho} \lambda_{2,t}) \mathcal{N}_t^{\frac{1}{\gamma}} e^{-\frac{\eta}{\gamma} t} = r\mathcal{G}, \end{aligned} \tag{5.36}$$

with the boundary condition,

$$\mathcal{G}(\mathcal{N}_T, \lambda_{2,T}, T) = 0.$$

Because a PDE of similar form has been solved by Wachter [2002], I take (5.32) and (5.33) as the guessed form of solution here. Substituting them into (5.36) and matching the coefficients of $\lambda_{2,t}^2$, $\lambda_{2,t}$ and the constant term produces a system of three differential equations very similar to (5.22). And their solution is (5.34). □ □

At first glance, it seems hard to understand why the differential equation (5.36) should have a solution in the integral form as in (5.32). This guessed solution, however, may follow naturally when utilizing the link between the intermediate consumption analysis and the terminal wealth analysis. Consider a series of auxiliary investors deriving utility from terminal wealth at time $i \in [0, T]$, and each investor is indexed by her fixed horizon date i. Suppose that investor i has initial wealth

$$W_{i,0} = W_0 \frac{E\left[\xi_i^{1-\frac{1}{\gamma}}\right] e^{-\frac{\eta}{\gamma} i}}{E\left[\int_0^T \xi_t^{1-\frac{1}{\gamma}} e^{-\frac{1}{\gamma}\eta t} dt\right]}. \tag{5.37}$$

Applying the terminal wealth analysis to investor i, we can write her optimal wealth at $t \in [0, i]$ as

$$\mathcal{W}_{i,t} := \mathcal{F}(\mathcal{N}_t, \lambda_{2,t}, t; i) = e^{-\frac{\eta}{\gamma} i} \mathcal{N}_t E \left[\mathcal{N}_i^{\frac{1}{\gamma}-1} \Big| \lambda_{2,t}, \mathcal{N}_t \right]$$

With the introduction of $\mathcal{F}(\mathcal{N}_t, \lambda_{2,t}, t; i)$, the optimal wealth of the investor with utility over consumption $\mathcal{G}(\mathcal{N}_t, \lambda_{2,t}, t)$ can be characterized as the sum of the optimal wealth of the auxiliary investors:

$$\mathcal{G}(\mathcal{N}_t, \lambda_{2,t}, t) = \int_t^T \mathcal{F}(\mathcal{N}_t, \lambda_{2,t}, t; s) ds. \tag{5.38}$$

The terminal wealth analysis can yield a solution of $\mathcal{F}$ that is similar to F given by (5.17). So the solution (5.32) follows immediately.

The derivation through a series of auxiliary investors has an interesting economic interpretation. $W_{i,0}$ given in (5.37) is the value at time zero of the optimal terminal wealth at time i. From (5.29) and (5.30), it follows that $W_{i,0}$ is the value at time zero of the optimal consumption event at period i for the investor concerned with intermediate consumption. The fraction at the right hand side of (5.37) is the ratio of period-i consumption to her life-time consumption in terms of the present value. Therefore, it is correct to think of the investor as holding separate accounts for each future consumption event, distributing her initial wealth into each account according to (5.37) to achieve the optimal consumption plan, and then investing each account so that the consumption needs are met.

5.4.2 Optimal portfolio and consumption policy

In the intermediate consumption case, the optimal financial strategy is characterized by the following theorem.

Theorem 5.4.2. *Suppose that in the Bond-Stock-Commodity economy, an investor seeks to maximize the expected utility over life-time consumption by choosing consumption and investment plans, as formalized in (5.13). Then the optimal consumption plan can be characterized by the following consumption-wealth ratio,*

$$\frac{c_t^\star}{\mathcal{W}_t} = \left[\int_t^T \mathcal{H}(\lambda_{2,t}, s-t) ds \right]^{-1}. \tag{5.39}$$

The optimal portfolio plan, denoted by $\boldsymbol{x}_t^\star := \begin{bmatrix} x_{S,t}^\star & x_{M,t}^\star \end{bmatrix}^\top$, *is*

$$\boldsymbol{x}_t^\star = \underbrace{\frac{1}{\gamma}\begin{bmatrix} \frac{\lambda_1}{\sigma_1} - \frac{\rho}{\sigma_1\bar{\rho}}\lambda_{2,t} \\ \frac{1}{\sigma_2\bar{\rho}}\lambda_{2,t} \end{bmatrix}}_{myopic\ part} + \underbrace{\frac{1}{\gamma}\begin{bmatrix} 0 \\ \frac{\beta}{\bar{\rho}}\frac{\int_t^T \mathcal{H}(\lambda_{2,t},s-t)[A_1(s-t)\lambda_{2,t}+A_2(s-t)]ds}{\int_t^T \mathcal{H}(\lambda_{2,t},s-t)ds} \end{bmatrix}}_{intertemporal\ part}. \tag{5.40}$$

Proof. The optimal consumption-wealth ratio follows from (5.30) and (5.32). The optimal portfolio plan can be obtained by equating the diffusion terms in (5.11) and (5.35), the two characterizations of $\mathcal{W}_t$. □ □

For the optimal consumption-wealth ratio, note that it changes with the commodity price, but not with the stock price.

The myopic allocation in the portfolio plan is the same as that in the terminal wealth case. It is a natural outcome when considering that in the analogy of the investor with utility over consumption to a series of investors concerned with terminal wealth, each of the auxiliary investors has identical myopic allocation.

The only new element arising from assuming intermediate consumption is contained in the intertemporal allocation to the commodity. Comparing (5.40) and (5.23), it is clear that the intertemporal allocation in the intermediate consumption case is a weighted average of that in the terminal wealth case, using $\mathcal{H}(\lambda_{2,t},\tau)$ as the weight. For $\mathcal{H}$, (5.39) implies

$$\frac{\mathcal{W}_t}{c_t^\star} = \int_t^T \mathcal{H}(\lambda_{2,t}, s-t)ds.$$

Hence, $\mathcal{H}(\lambda_{2,t},\tau)$ can be interpreted as the time-t value of future consumption in τ periods normalized by the optimal consumption rate at time-t. In all, the intertemporal allocation assuming intermediate consumption is an average of those assuming terminal wealth, and the average is weighted by the value of future consumption in each period.

5.4.3 Welfare analysis

When people are concerned with interim consumption, the utility loss of leaving out the commodity is as follows.

Proposition 5.4.3. *Suppose that in the Bond-Stock-Commodity economy, an investor is concerned with maximizing the expected utility over life-time consumption as described in (5.13). If the commodity is excluded in consumption and portfolio decisions, then* $\mathcal{L}$

percent of extra initial wealth is needed to achieve the same expected utility level as is obtained by following the optimal strategy given by Theorem 5.4.2, and

$$\frac{\mathcal{L}}{100} = \left[\frac{\omega \int_0^T \mathcal{H}(\lambda_{2,0}, s)ds}{1 - e^{-\omega T}}\right]^{\frac{\gamma}{1-\gamma}} - 1. \tag{5.41}$$

where

$$\omega := \frac{\eta - r(1-\gamma)}{\gamma} - \frac{1}{2}\frac{1-\gamma}{\gamma^2}\lambda_1^2.$$

Proof. By reasoning similar to that in the terminal wealth case, the indirect utility function assuming intermediate consumption, denoted by $\mathcal{J}(\mathcal{W}_t, \lambda_{2,t}, t)$, is given by

$$\mathcal{J}(\mathcal{W}_t, \lambda_{2,t}, t) = \frac{\mathcal{W}_t^{1-\gamma}}{1-\gamma} e^{-\eta t} \left[\int_t^T \mathcal{H}(\lambda_{2,t}, s-t)ds\right]^{\gamma}.$$

In the standard Black-Scholes economy after dropping the commodity, the corresponding indirect utility is

$$\mathcal{J}^{BS}(\mathcal{W}_t, t) = \frac{\mathcal{W}_t^{1-\gamma}}{1-\gamma}\left[\frac{1}{\omega}\left(1 - e^{-\omega(T-t)}\right)\right]^{\gamma},$$

Then $\mathcal{L}$ as given by (5.41) follows from its definition. □ □

5.5 Calibration and discussion

In this section, I shall first estimate the commodity futures price model. Then some representative numerical illustrations and discussions will be presented.

5.5.1 Estimation of the commodity futures model

The parameters that characterize the asset class of commodities are estimated using the GSCI Commodities Index futures prices. The GSCI index underlying the futures contract tracks the price levels of major commodities, and the futures contract can be viewed as being written on a basket of commodities. Therefore, the GSCI Commodities Index is taken to be the non-tradable spot commodity price, and the GSCI Commodities Index futures prices are the tradable futures prices.

The estimation is carried out in two steps. In the first, the three parameters that characterize the spot commodity price, θ, μ_2, and σ_2, are estimated. From (5.3), the logarithm of the spot commodity price follows a first-order autoregressive process, and the maximum likelihood method is used to get the estimates (Table 5.1). I use the monthly data of GSCI Commodity Index from December 1969, the start date of the index, to

November 2008. The data are deflated by the US CPI-U index, for the reason that the asset prices are assumed to be measured in real terms in the Bond-Stock-Commodity economy. The deflated data are normalized in such a way that the value was 100 in July 1992 when Chicago Mercantile Exchange (CME) introduced futures on this index.

The second step is to estimate α and β, which specify the risk premium in the futures market (5.4). The futures price (5.5) can, in log form, be rewritten as

$$\ln F_t(\mathcal{T}) = e^{-\widetilde{\theta}\mathcal{T}} m_t + (1 - e^{-\widetilde{\theta}\mathcal{T}})\frac{\widetilde{\mu}_2}{\widetilde{\theta}} + \frac{\sigma_2^2}{4\widetilde{\theta}}(1 - e^{-2\widetilde{\theta}\mathcal{T}}), \tag{5.42}$$

where $\mathcal{T}$ is the time to maturity of the futures contract. From Ito's formula and (5.3) it follows

$$d\ln F_t(\mathcal{T}) = \theta e^{-\widetilde{\theta}\mathcal{T}} \left(\mu_2 - \frac{\sigma_2^2}{2\theta} - m_t\right) dt + \sigma_2 e^{-\widetilde{\theta}\mathcal{T}} dZ_{M,t}. \tag{5.43}$$

Inserting (5.42) into (5.43) leads to

$$\begin{aligned} d\ln F_t(\mathcal{T}) = \theta \left[e^{-\widetilde{\theta}\mathcal{T}} \left(\mu_2 - \frac{\sigma_2^2}{2\theta}\right) + \left(1 - e^{-\widetilde{\theta}\mathcal{T}}\right) \frac{\widetilde{\mu}_2}{\widetilde{\theta}} + \frac{\sigma_2^2}{4\theta} \left(1 - e^{-2\widetilde{\theta}\mathcal{T}}\right) - \ln F_t(\mathcal{T})\right] dt \\ + \sigma_2 e^{-2\widetilde{\theta}\mathcal{T}} dZ_{M,t}. \end{aligned} \tag{5.44}$$

From (5.44) it is clear that the futures price in log form with a constant time-to-maturity $\mathcal{T}$, $\ln F_t(\mathcal{T})$, also follows a first-order autoregressive process, and the maximum likelihood method can be used to estimate α and β (through estimating $\widetilde{\theta}$ and $\widetilde{\mu}$).

Data of the GSCI Commodities Index futures prices are used to estimate (5.44). In view of the liquidity of futures contracts, the first three nearby futures contracts are used in the estimation. In particular, the data, obtained from Bloomberg, consist of the monthly observations of the prices of these three contracts at the end of each month from July 1992 to November 2008. Since the futures trading terminates on the eleventh business day of the contract month, the time to maturity for each of these three contracts does not change with the observations. As with the underlying index, the futures prices are deflated by the CPI-U index, and normalized correspondingly.

The estimates of α and β are presented in Table 5.1. The estimate of β is significantly negative, suggesting that the risk premium of the commodity futures is decreasing in the spot commodity price.

For the purpose of numerical illustration, I assume the other parameter values as given in Table 5.1. The parameter values for the stock and the riskless bond, μ_1, σ_1, and r, are taken to be consistent with many empirical studies, e.g. Campbell [2003].

The value of the correlation coefficient is chosen to be -0.10, a level corresponding to the finding in many studies that the commodity and stock returns have a moderate negative correlation.

5.5.2 Optimal strategy and utility of commodity investment

For the parameter values as given in Table 5.1, optimal financial strategies and utility losses are determined. The focus of this numerical exercise is on the influence of the commodity price, of the horizon of the investor, and of the risk aversion of the investor. From the property of Ornstein-Uhlenbeck processes, it follows that the logarithm of commodity price is asymptotically stationary, and the asymptotically stationary distribution is normal: $\Phi\left(\mu_2 - \frac{\sigma_2^2}{2\theta}, \frac{\sigma_2^2}{2\theta}\right)$. To have an intuitive idea of the level of current commodity price, I shall locate it with respect to this asymptotical distribution.

Figure 5.2 shows the optimal strategies and utility losses for a range of current commodity prices from 66 through 251. With respect to the asymptotically stationary distribution of the commodity price, this commodity price range corresponds to the one from the 5th-percentile to the 95th-percentile. For this range of commodity prices, the expected return on the representative portfolio of commodity futures varies from 9.42% to -3.62%. In this example, the optimal stock weight is decreasing with the spot commodity price, owing to the assumed negative correlation coefficient between the two driving Brownian motions, and the decreasing relationship between the risk premium in the commodity futures market and the spot commodity price. Notably, the commodity allocation varies considerably with the changes of the commodity price. The commodity allocation decreases from a long position of almost 70% for the commodity price at 5th-percentile, to a short position of around 25% for the commodity price at 95th-percentile. In the optimal portfolio strategy, there is substantial market timing.

We have learned that the utility losses of leaving out the commodity depend on the

Estimated parameter values					Assumed parameter values				
$\hat{\theta}$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\alpha}$	$\hat{\beta}$	μ_1	σ_1	r	ρ	η
0.120	5.023	0.199	2.335	-0.465	0.080	0.150	0.010	-0.100	0.010
(0.082)	(0.066)	(0.012)	(0.518)	(0.186)					

Table 5.1: **The parameter values used for the numerical exercise** The parameters characterizing the asset class of commodities are estimated, and the standard errors are in parenthesis. The other parameter values are taken to be consistent with many existing studies.

current commodity price (Propositions 5.3.3 and 5.4.3). In this numerical example, the current commodity price has a significant impact on the magnitude of utility loss. For the given range of commodity price, the utility loss varies from 4% to 22% in the terminal wealth case, and from 2% to 10% in the intermediate consumption case. It suggests that excluding commodities in financial decisions is much more costly when commodity prices are very low or very high than when they are moderate.

Figure 5.3 shows that the utility loss of excluding the commodity is increasing in the horizon, T, and decreasing in the degree of risk aversion, γ. The decreasing relation to the degree of risk aversion is perhaps what one would expect, considering that risky assets, including commodities, account for a lower portfolio weight for more risk averse people as they would invest more in the riskless asset. The increasing relation to the horizon also should come as no surprise if considering that a longer period of time enables one to benefit more from the additional investment opportunities brought by commodities. These intuitions make it tempting to conjecture that these relations hold independent of

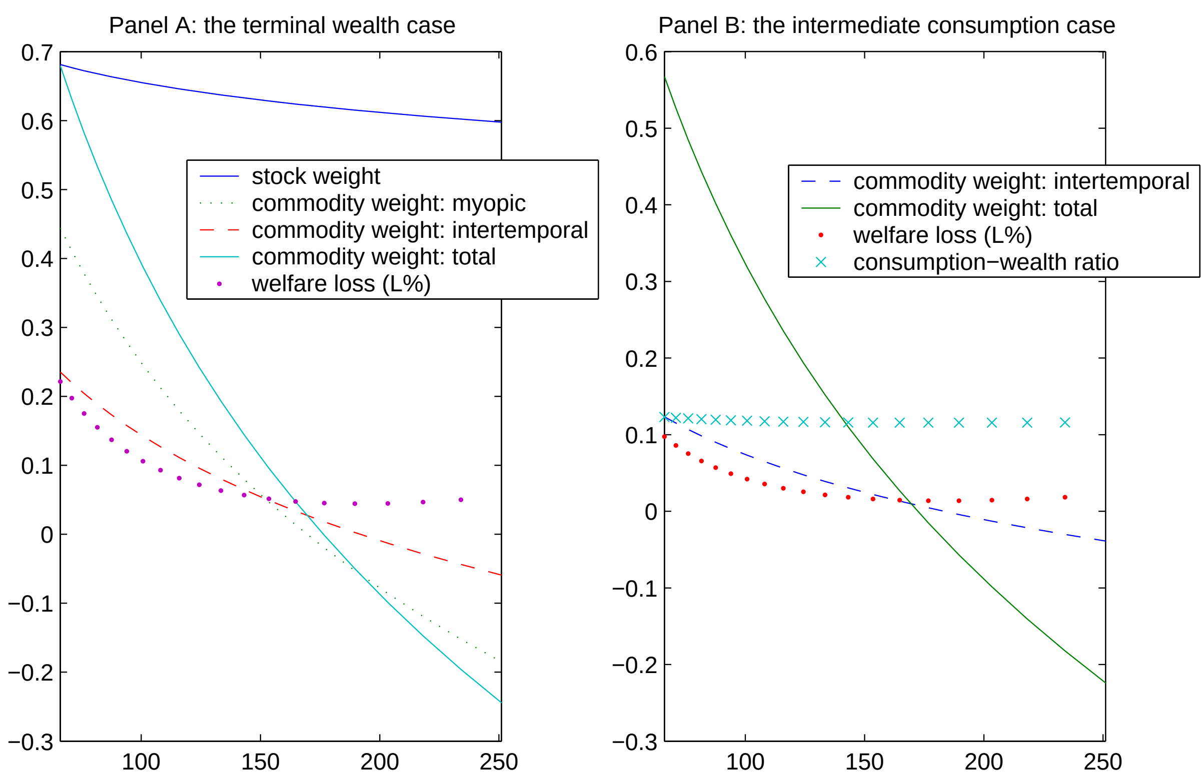

Figure 5.2: **Optimal policy and the commodity price** This figure shows how the optimal strategy and the utility loss changes with the current commodity price in the terminal wealth case (Panel A), and in the intermediate consumption case (Panel B). It is assumed that $T = 10$, and $\gamma = 5$. The stock allocation and the myopic allocation to the commodity are the same in both cases, and they are not repeated in Panel B for ease of reading.

the chosen parameter values.

Figure 5.3 also shows that other things being equal, excluding the commodity is less costly in the intermediate consumption case than in the terminal wealth case. Other things being equal, the utility cost of excluding the commodity is increasing in the horizon, then a plausible explanation for the lower utility cost in the intermediate consumption case could be that the "effective" horizon of the investor assuming intermediate consumption is shorter than that assuming terminal wealth. The aforementioned analogy to coupon-bearing and zero-coupon bonds helps to understand the explanation. It is well known that a coupon-bearing bond has a shorter effective horizon than its zero-coupon counterpart. By analogy, the investor assuming intermediate consumption should have a shorter effective horizon than her counterpart assuming terminal wealth, and hence suffers a lower utility cost.

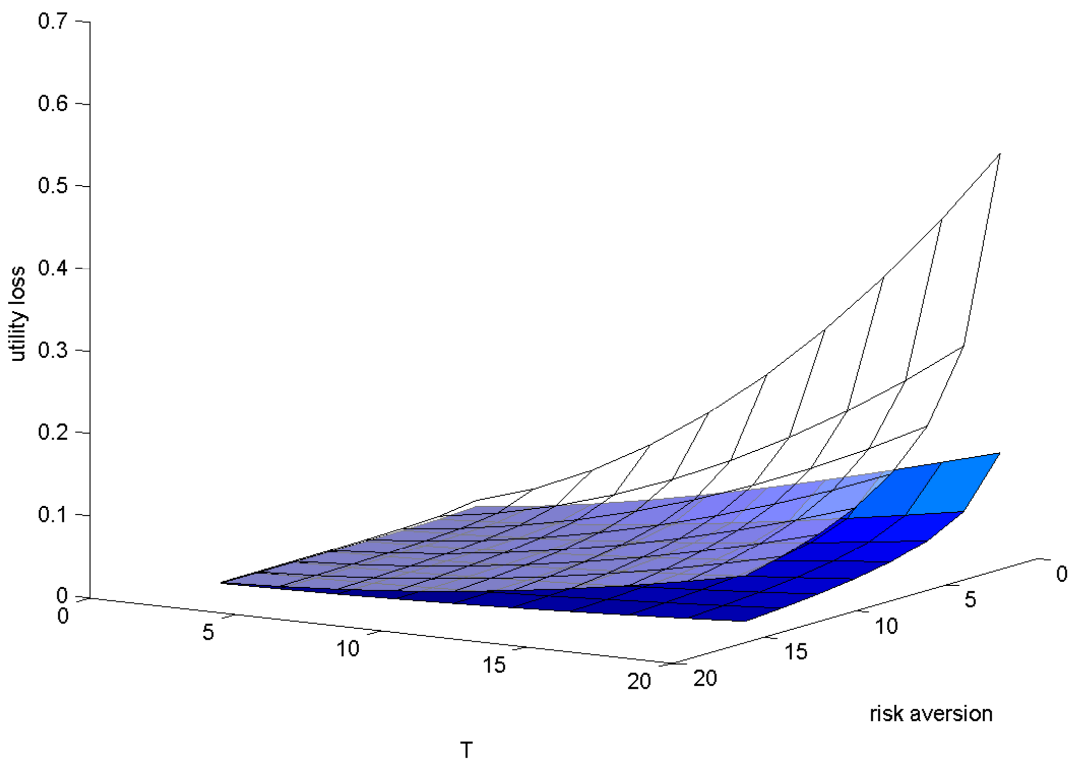

Figure 5.3: **The effect on utility loss of investment horizon and risk aversion** This figure shows how the utility loss changes with the investment horizon and the degree of risk aversion in the terminal wealth case and the intermediate consumption case. The spot commodity price is assumed to be 129, the median value of the asymptotic distribution.

The finding that the utility loss is decreasing in risk aversion is different from that of Anson [1999], who concludes that the more risk-averse the investor, the higher the utility of investing in commodity futures. A plausible explanation of this difference lies in the different framework of analysis: Anson's study is based on a one-period mean-variance framework with only risky assets, while our conclusion comes from a dynamic framework where the riskless asset is available. Anson [1999] shows that the allocation to commodity futures is monotonically increasing with the degree of risk aversion. In comparison, this numerical exercise shows the effect of risk aversion on commodity allocation is not monotonic (Figure 5.4). We can distinguish two cases: (i) when the spot commodity price is relatively low, the commodity weight is decreasing with risk aversion; and (ii) when the spot commodity price is relatively high, the commodity weight is increasing with risk aversion.

Now turn to the intertemporal allocation to the commodity. Figure 5.5 shows how the intertemporal commodity allocation changes with the investment horizon and the degree

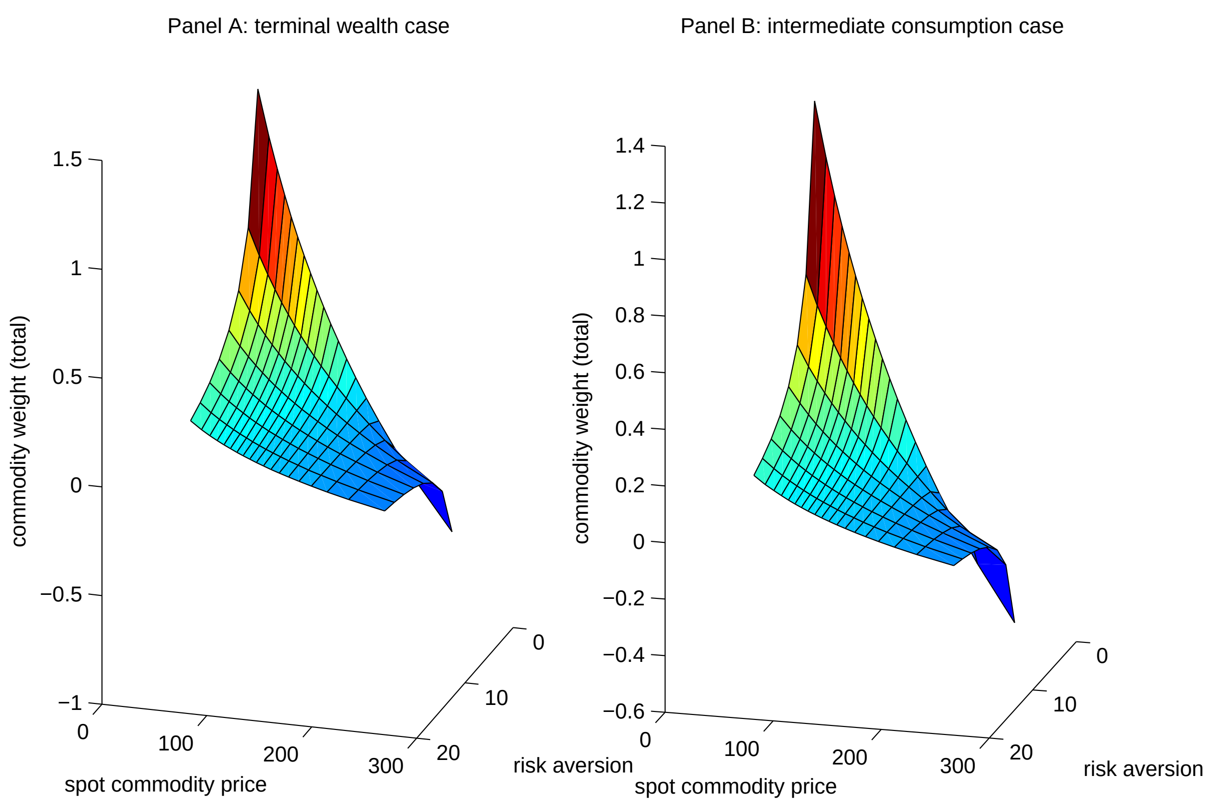

Figure 5.4: **The effect on the optimal commodity weight of the spot commodity price and risk aversion** This figure shows how the total commodity weight changes with the spot commodity price and the degree of risk aversion in the terminal wealth case and the intermediate consumption case. It is assumed that $T = 10$.

of risk aversion. Two interesting points stand out from this figure: (i) in both terminal wealth and intermediate consumption cases, the commodity allocation increases with the horizon of the investor; and (ii) the investor with utility over terminal wealth allocates more to the commodity than the investor concerned with intermediate consumption. Actually, this figure illustrates the following proposition.

Proposition 5.5.1. *If $\lambda_{2,t}$ and λ^* are positive, then other things being equal, the optimal commodity allocation increases with the horizon in both the terminal wealth case and the intermediate consumption case. Moreover, the optimal intertemporal allocation to the commodity is greater in the terminal wealth case than in the intermediate consumption case.*

Proof. When λ^* is positive, then $A_1(\tau)$, $A_2(\tau)$ and their derivatives are all negative (see

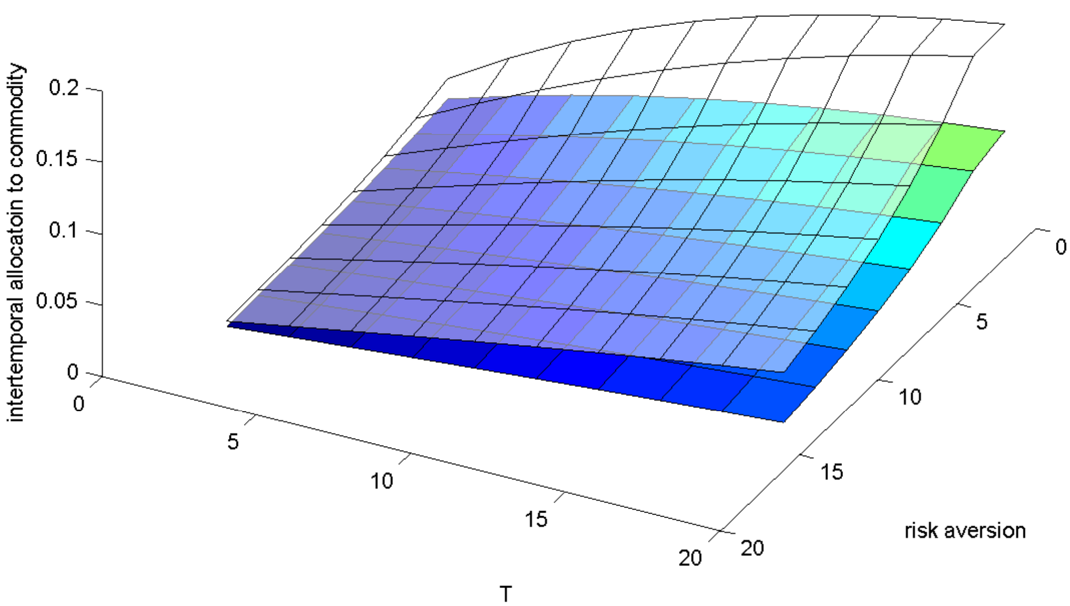

Figure 5.5: **The effect of horizon and risk aversion on the intertemporal commodity allocation** This figure shows how the intertemporal commodity allocation changes with the horizon and the degree of risk aversion in the terminal wealth case and the intermediate consumption case. The spot commodity price is assumed to be 129, the median value of the asymptotic distribution.

Appendix 5.A). This implies that

$$
\begin{aligned}
-A_1(T-t) &> -\frac{\int_t^T \mathcal{H}(\lambda_{2,t}, s-t) A_1(s-t) ds}{\int_t^T \mathcal{H}(\lambda_{2,t}, s-t) ds} \\
-A_2(T-t) &> -\frac{\int_t^T \mathcal{H}(\lambda_{2,t}, s-t) A_2(s-t) ds}{\int_t^T \mathcal{H}(\lambda_{2,t}, s-t) ds}.
\end{aligned}
$$

From (5.23) and (5.40) and condition that $\lambda_{2,t}$ is positive, the statement about the intertemporal allocation follows. □ □

Note that given the parameter values and the current commodity prices used in Figure 5.5, λ^* and $\lambda_{2,t}$ are positive.

In sharp contrast to the significant utility of commodity investment in the dynamic framework, the commodity market in this example is of little utility for a one-period mean-variance investor. With these parameter values, the unconditional Sharpe ratio in the commodity market, $E[\lambda_{M,t}]$, is 7.6%. In comparison, the Sharpe ratio in the stock market is 47%. Given an empirically reasonable value of the correlation coefficient, the allocation to and utility of the commodity investment in a mean-variance optimized portfolio would be of little practical relevance. Therefore, investing in this asset class makes little sense in the one-period mean-variance optimization framework. But under the dynamic framework used in this study, commodity allocation may be needed in considerable magnitude, and allocation to the asset class of commodities may yield a consequential welfare improvement. The difference in the perception of the asset class of commodities stems from the difference in the research frameworks.

The large utility of commodity investment illustrated in this example relies on exploiting the return predictability in the commodity market. In this study, however, I abstract from robustness issues, especially uncertainty about the predictive relationship, which would affect the optimal strategy and the utility improvement arising from commodity investment. Further research is needed to address these issues.

5.6 Conclusion

Given the rapid growth of commodity markets as a major alternative asset class, this study has aimed at examining this asset class in the framework of dynamic asset allocation à la Merton [1969]. In this study, the risk premium in the commodity market is assumed to be time-varying and dependent on the non-tradable spot commodity price, while the spot commodity price has the property of mean reversion. This characterization of the commodity market, an extended version of the single-factor model in Schwartz [1997], is consistent with the empirical finding in this chapter that the spot commodity price has significantly negative effect on the risk premium. As such, this study underscores the implications for portfolio and consumption decisions of the time-varying risk premia of commodity investment arising from the mean-reverting nature of commodity prices.

I have derived the optimal dynamic strategies for an investor who is concerned either with terminal wealth or with intermediate consumption. Closed-form expressions were obtained for the optimal strategies and the utility losses of excluding the commodity in financial decisions. In the optimal policy, the allocation to commodities is made for myopic purposes, as well as for intertemporal hedging purposes. The optimal allocation to the stock is solely myopic, and dependent on the spot commodity price.

Based on some representative parameter values, I presented a numerical example for the optimal strategy and the utility cost of leaving out the commodity. In this example, as long as the risk premium in the commodity market is positive, the optimal commodity allocation has a horizon effect—the longer horizon, the more allocation to the commodity. The example also shows that if the investor is relatively less risk averse, and concerned with financial decisions over a longer time period, then the commodity market brings a greater utility improvement. It suggests that in a long-term financial plan, like that of saving for retirement, it is rather costly to exclude commodities from financial decision making.

In this example, the significant utility of commodity investment is established from the dynamic perspective on financial decisions, while an investor from the static mean-variance perspective would find little value in commodity investment. In the static mean-variance framework, an appreciably positive risk premium in commodity markets is needed for investors to include commodities in mainstream portfolios as an appealing asset class. When dynamic investment strategies are allowed, however, considerable utility of commodity investment may come from exploiting the return predictability in commodity markets.

5.A Appendix

The portfolio strategy of the representative commodity portfolio $\mathcal{M}_t$

Applying Ito's formula to the futures price (5.5) yields

$$dF_t(T) = F_t(T)e^{-\widetilde{\theta}(T-t)}\sigma_2\lambda_{M,t}dt + F_t(T)e^{-\widetilde{\theta}(T-t)}\sigma_2 dZ_{M,t}.$$

Consider a self-financing portfolio $\mathcal{M}_t$ consisting of ψ_t units of the riskless bond and ϕ_t units of the futures contract. Because the futures contract is constantly resettled to have zero value and the portfolio is self-financing, the portfolio $\mathcal{M}_t$ satisfies

$$\mathcal{M}_t = \psi_t B_t, \quad \text{and} \quad d\mathcal{M}_t = \psi_t dB_t + \phi_t dF_t(T).$$

If the units of the futures contract $\phi_t = \frac{\mathcal{M}_t}{F_t(T)}e^{\widetilde{\theta}(T-t)}$, then

$$d\mathcal{M}_t = \mathcal{M}_t\left(r + \sigma_2\lambda_{M,t}\right)dt + \mathcal{M}_t\sigma_2 dZ_{M,t}.$$

Note that the above equation only holds for $t \leq T$. To overcome this restriction, we can consider rolling over the futures contract. Namely, the holding of futures contract is rolled over to keep the time to maturity of the held contract equal to a constant ℓ, and the units of futures contract $\phi_t = \frac{\mathcal{M}_t}{F_t(T)}e^{\widetilde{\theta}\ell}$. Under this portfolio strategy, the notional value of the futures holding is

$$\phi_t F_t(t+\ell) = \mathcal{M}_t e^{\widetilde{\theta}\ell}.$$

Validity of the optimal policy

This appendix shows the validity of the optimal strategies that have been derived in the terminal wealth case and in the interim consumption case. In addition to the assumption that $E(\xi_T^{-1})$ is finite, the following two conditions need to be verified [Cox and Huang, 1989, Theorem 2.2]: (i) the Lagrange multipliers k and $\mathcal{K}$ are positive and finite; and (ii) F and $\mathcal{G}$ have sufficient continuous differentiability.

The Lagrange multipliers k and $\mathcal{K}$ are positive and finite From (5.10), it follows that

$$\begin{aligned}\xi_T &= \exp\left[-\left(r + \frac{1}{2}\lambda_1^2\right)T - \lambda_1 Z_{1,T} - \frac{1}{2}\int_0^T \lambda_{2,t}^2 dt - \int_0^T \lambda_{2,t} dZ_{2,t}\right]\\ &\leq \exp\left[-\left(r + \frac{1}{2}\lambda_1^2\right)T - \lambda_1 Z_{1,T} - \int_0^T \lambda_{2,t} dZ_{2,t}\right].\end{aligned}$$

It can be seen that ξ_T is positive. Moreover, because the right hand side is lognormal, $E[\xi_T]$ is bounded by the lognormal variable's expectation, which is finite and a continuous function of time. The two Lagrange multipliers are given by

$$\begin{aligned} k &= W_0^{-\gamma}\left(E\left[\xi_T^{1-1/\gamma}\right]\right)^{\gamma} \\ \mathcal{K} &= W_0^{-\gamma}\left(E\left[\int_0^T \xi_t^{1-\frac{1}{\gamma}} e^{-\frac{1}{\gamma}\eta t} dt\right]\right)^{\gamma}. \end{aligned}$$

Their positiveness follows immediately, and their finiteness follows from Jensen's inequality and the assumption that $\gamma > 1$.

Continuous differentiability of F and $\mathcal{G}$ From (5.17) and (5.32), the differentiability condition holds inasmuch as $A_1(\tau)$, $A_2(\tau)$, $A_3(\tau)$ and $\mathcal{A}_3(\tau)$ are continuously differentiable. From (5.19) and (5.34), it follows that they are indeed continuously differentiable (for the continuous differentiability of $A_1(\tau)$ and $A_1(\tau)$, see (5.45) below).

Properties of $A_1(\tau)$ and $A_2(\tau)$

Define

$$B(\tau) := 2q - (q + a_1)\left(1 - e^{-q\tau}\right).$$

Because $q > 0$, and $q > a_1$, $B(\tau) > 0$. From the assumption that $\gamma > 1$, it follows that

$$\begin{aligned} A_1(\tau) &< 0, \\ \operatorname{sign}\big(A_2(\tau)\big) &= -\operatorname{sign}(\lambda^*). \end{aligned}$$

The derivatives of $A_1(\tau)$ and $A_2(\tau)$ are

$$\begin{aligned} A_1'(\tau) &= \frac{1-\gamma}{\gamma}\frac{4q^2 e^{-q\tau}}{B(\tau)^2} \\ A_2'(\tau) &= \frac{1-\gamma}{\gamma}\frac{4\lambda^*\left(e^{-q\tau/2} - e^{-q\tau}\right)\left[q - a_1 + (q + a_1)e^{-q\tau/2}\right]}{B(\tau)^2} \end{aligned} \tag{5.45}$$

So $A_1'(\tau)$ and $A_2'(\tau)$ are continuous, and their signs are

$$\begin{aligned} A_1'(\tau) &< 0, \\ \operatorname{sign}\left(A_2'(\tau)\right) &= -\operatorname{sign}(\lambda^*). \end{aligned}$$

References

K. Aase and S. Persson. Valuation of the minimum guaranteed return embedded in life insurance products. *Journal of Risk and Insurance*, 64:599–617, 1997.

Y. Ait-Sahalia and M. Brandt. Variable selection for portfolio choice. *Journal of Finance*, 56:1297–1351, 2001.

H. Andersson. Are commodity prices mean reverting? *Applied Financial Economics*, 17:769–783, 2007.

M.J.P. Anson. Maximizing utility with commodity futures diversification. *Journal of Portfolio Management*, Summer:86–94, 1999.

L. Ballotta, S. Haberman, and N. Wang. Guarantees in with-profit and unitized with-profit life insurance contract: fair valuation problem in presence of the default option. *Journal of Risk and Insurance*, 73:97–121, 2006.

N. Barberis. Investing for the long run when returns are predictable. *Journal of Finance*, 55:225–264, 2000.

N. Barberis, M. Huang, and T. Santos. Prospect theory and asset prices. *Quarterly Journal of Economics*, 116:1–53, 2001.

D. Bauer, R. Kiesel, A. Kling, and J. Ruß. Risk-neutral valuation of participating life insurance contract. *Insurance: Mathematics and Economics*, 39:171–183, 2006.

S. Benartzi and R. Thaler. Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, 110:73–92, 1995.

A. Berkelaar, R. Kouwenberg, and T. Post. Portfolio choice under loss aversions. *Review of Economics and Statistics*, 64:973–986, 2004.

H. Bessembinder and K. Chan. Time-varying risk premia and forecastable returns in futures markets. *Journal of Financial Economics*, 32:169–193, 1992.

H. Bessembinder, J.F. Coughenour, P.J. Seguin, and M.M. Smoller. Mean reversion in equilibrium asset prices: evidence from the futures term structure. *Journal of Finance*, 50:361–375, 1995.

J. Bikker and P. Vlaar. Conditional indexation in defined benefit pension plans. *The Geneva Papers on Risk and Insurance Issues and Practice*, 32:494–515, 2007.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

Z. Bodie and V. Rosansky. Risk and return in commodity futures. *Financial Analysts Journal*, 36:3–14, 1980.

M. J. Brennan. The role of learning in dynamic portfolio decisions. *European Economic Review*, 1:295–306, 1998.

S. Browne. The risk and rewards of minimizing shortfall probability. *Journal of Portfolio Management*, 25:76–85, 1999.

J. Y. Campbell. Consumption-based asset pricing. In G. Constantinides, M. Harris, and R. Stultz, editors, *Handbook of the Economics of Finance*, volume 1B, chapter 13, pages 803–887. North-Holland, Amsterdam, 2003.

J. Y. Campbell and J. H. Cochrane. By force of habit: a consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107: 205–251, 1999.

J. H. Cochrane. *Asset Pricing*. Princeton University Press, Princeton, New Jersey, revised edition, 2005.

J. Cox and C. Huang. Optimum consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49:33–83, 1989.

J. Cox and C. Huang. A variational problem arising in financial economics. *Journal of Mathematical Economics*, 20:465–487, 1991.

F. de Jong. Pension fund investment and the valuation of liabilities under conditional indexation. *Insurance: Mathematics and Economics*, 42:1–13, 2008.

J. B. Detemple. Asset pricing in an economy with incomplete information. *Journal of Finance*, 41:383–391, 1986.

N. Dokuchaev. Mean-reverting market model: speculative opportunities and non-arbitrage. *Applied Mathematical Finance*, 14:319–337, 2007.

M. U. Dothan and D. Feldman. Equilibrium interest rates and multiperiod bonds in a partially observed economy. *Journal of Finance*, 41:369–382, 1986.

E. Doyle, J. Hill, and I. Jack. Growth in commodity investment: risks and challenges for commodity market participants. Markets Infrastructure Department, Fianncial Services Authority, March 2007.

D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey, third edition, 2001.

C. B. Erb and C. Harvey. The strategic and tactical value of commodity futures. *Financial Analysts Journal*, 62:69–97, 2006.

E. F. Fama and K. R. French. Business cycles and the behavior of metals prices. *Journal of Finance*, 43:1075–1093, 1988.

E. F. Fama and K. R. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 29:23–49, 1989.

W. H. Fleming and R. W. Rishel. *Deterministic and Stochastic Optimal Control*. Springer-Verlag, New York, 1975.

I. Friend and M. E. Blume. The demand for risky assets. *American Economic Review*, 65:900–922, 1975.

J. Gabillon. Analysing the forward curve. In *Managing Energy Price Risk*. Risk Publications, London, 1995.

N. Gatzert and A. Kling. Analysis of participating life insurance contract: a unification approach. *Journal of Risk and Insurance*, 74:547–570, 2007.

H. Geman. *Commodities and Commodity Derivatives: Modelling and Pricing for Agriculturals, Metals and Energy*. John Wiley & Sons, West Sussex, England, 2005.

G. Gennotte. Optimal portfolio choice under incomplete information. *Journal of Finance*, 41:733–746, 1986.

F. J. Gomes. Portfolio choice and trading volume with loss-averse investors. *Journal of Business*, 78:3675–706, 2005.

G. Gordon and G. Rouwenhorst. A note on Erb and Harvey (2005). Yale School of Management, 2005.

G. Gordon and G. Rouwenhorst. Facts and fantasies about commodity futures. *Financial Analysts Journal*, 62:47–68, 2006.

R. J. Greer. Conservative commodities: a key inflation hedge. *Journal of Portfolio Management*, 21:60–77, 1978.

A. Grosen and P. Jørgensen. Life insurance liabilities at market value. *Journal of Risk and Insurance*, 69:63–91, 2002.

R.P.M.M. Hoevenaars, R.D.J. Molenaar, P.C. Schotman, and T.B.M. Steenkamp. Strategic asset allocation with liabilities: beyond stocks and bonds. *Journal of Economic Dynamics and Control*, 32:2939–2970, 2008.

J. E. Ingersoll. *Theory of Financial Decision Making*. Rowman & Littlefield, Totowa, N.J., 1987.

G. R. Jensen, R. R. Johnson, and J. M. Mercer. Efficient use of commodity futures in diversified portfolios. *Journal of Futures Markets*, 20(5):88–98, 2000.

G. R. Jensen, R. R. Johnson, and J. M. Mercer. Tactical asset allocation and commodity futures. *Journal of Portfolio Management*, 28(4):100–111, 2002.

D. Kahneman and A. Tversky. Prospect theory: an analysis of decision under risk. *Econometrica*, 47:263–290, 1979.

I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, New York, 1998.

H. M. Kat and R. C. A. Oomen. What every investor should know about commodities Part I: univariate return analysis. *Journal of Investment Management*, 5(1):4–28, 2007.

T. S. Kim and E. Omberg. Dynamic nonmyopic portfolio behavior. *Review of Financial Studies*, 9:141–161, 1996.

T. Kleinow and M. Willder. The effect of management discretion on hedging and fair valuation of participating policies with maturity guarantees. *Insurance: Mathematics and Economics*, 40:445–458, 2007.

T. Kocken. Constructing sustainable pensions. *Life & Pensions*, July-August:35–39, 2007.

T. P. Kocken. *Curious Contracts: Pension Fund Redesign for the Future*. PhD thesis, Vrije Universiteit, 2006.

J. M. Kreps and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stocastic Processes and Their Applications*, 11:215–260, 1981.

R. S. Liptser and A. N. Shiryayev. *Statistics of Random Processes II: Applications*. Springer-Verlag, New York, 1978.

A. Lusardi and O. Mitchell. Financial literary and planning: implication for retirement wellbeing. Dartmouth College and University of Pennsylvania, January 2006.

J. G. March and Z. Shapira. Managerial perspectives on risk and risk taking. *Management Science*, 33:1404–1419, 1987.

H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

R. C. Merton. Observations on innovation in pension fund management in the impending future. *PREA Quarterly*, pages 61–67, 2006.

R. C. Merton. Lifetime portfolio selection under uncertainty: the continuous-time case. *Review of Economics and Statistics*, 51:247–257, 1969.

R. C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3:373–413, 1971.

R. C. Merton. Theory of rational option pricing. *Bell Journal of Economics*, 4:141–183, 1973.

J. Miffre and G. Rallis. Momentum strategies in commodity futures markets. *Journal of Banking and Finance*, 31:1863–1886, 2007.

P. Mongars and C. Marchal-Dombrat. Commodities: an asset class in their own right? *Banque de France: Financial Stability Review*, 9, December 2006.

T. E. Nijman and R. S. J. Koijen. Valuation and risk management of inflation-sensitive pension rights. In N. E. Kortleve, T. E. Nijman, and E. P. M. Ponds, editors, *Fair Value and Pension Fund Management*. Elsevier, Amsterdam, 2006.

T. E. Nijman and L. A. P. Swinkels. Strategic and tactical allocation to commodities for retirement savings schemes. Tilburg University, 2007.

R. Pindyck. The dynamics of commodity spot and futures markets: a primer. *Energy Journal*, 22:1–29, 2001.

R. Pindyck. Risk aversion and the determinants of stock market behavior. *Review of Economics and Statistics*, 70:183–190, 1988.

J. M. Poterba and L. H. Summers. Mean reversion in stock prices: evidence and implications. *Journal of Financial Economics*, 22:27–59, 1988.

P. A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51:239–246, 1969.

E. S. Schwartz. The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, 52:923–973, 1997.

U. Segal and A. Spivak. First order versus second order risk aversion. *Journal of Economic Theory*, 51:111–125, 1990.

W. F. Sharpe. Morningstar's risk-adjusted ratings. *Financial Analysts Journal*, 55: 21–33, 1998.

K. Sullivan and T. Kida. The effect of multiple reference points and prior gains and losses on managers' risky decision making. *Organizational Behavior and Human Decision Process*, 64:76–83, 1995.

S. M. Sundaresan. Intertemporally dependent preferences and the volatility of consumption and wealth. *Review of Financial Studies*, 2, 1989.

G. G. Szpiro. Measuring risk aversion: an alternative approach. *Review of Economic and Statistics*, 68:156–159, 1986.

H. Till and J. Eagleeye, editors. *Intelligent Commodity Investing: New Strategies and Practical Insights for Informed Decision Making*. Risk Books, London, 2007.

A. Tversky and D. Kahneman. Decisions and the psychology of choice. *Science*, 211: 453–458, 1981.

A. Tversky and D. Kahneman. Loss aversion in riskless choice: a reference-dependent model. *Quarterly Journal of Economics*, 106:1039–1061, 1992.

M. van Rooij, A. Lusardi, and R. Alessie. Financial literacy and stock market participation. De Nederlandsche Bank, mimeo, 2006.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey, 1944.

E. B. Vrugt, R. Bauer, R. Molenaar, and T. Steenkamp. Dynamic commodity timing strategies. In H. Till and J. Eagleeye, editors, *Intelligent Commodity Investing: New Strategies and Practical Insights for Informed Decision Making*. Risk Books, London, 2007.

J. A. Wachter. Portfolio and consumption decisions under mean-reverting returns: an exact solution for complete markets. *Journal of Financial and Quantitative Analysis*, 37:63 – 91, 2002.

C. Wang and M. Yu. Trading activity and price reversals in futures markets. *Journal of Banking and Finance*, 28:1337–1361, 2004.

Y. Xia. Learning about predictability: the effect of parameter uncertainty on dynamic asset allocation. *Journal of Finance*, 56:205–246, 2001.